

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2001, 00:02:02 ; Search time 66.17 Seconds
(without alignments)
372.057 Million cell updates/sec

Title: US-09-331-723A-4
 Perfect score: 141
 Sequence: 1 gccgcgcagggccctggctc.....agctgcacccgcgcacgcag 141

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 302621 seqs, 87301344 residues

Total number of hits satisfying chosen parameters: 605242

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database : Issued_Patents_NA:*
1: /cgn2_6/pdata/1/lna/5A_COMB.seq:*
2: /cgn2_6/pdata/1/lna/5B_COMB.seq:*
3: /cgn2_6/pdata/1/lna/6A_COMB.seq:*
4: /cgn2_6/pdata/1/lna/6B_COMB.seq:*
5: /cgn2_6/pdata/1/lna/PCRTS_COMB.seq:*
6: /cgn2_6/pdata/1/lna/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	%		Length	DB	ID	Description
			Match					
C	1	139.4	98.9	3381	4	US-09-009-119-1	Sequence 1, Appli	
C	2	136.2	96.6	3383	5	PCR-US95-09098-1	Sequence 1, Appli	
C	3	38.8	27.5	2900	1	US-08-034-650-9	Sequence 9, Appli	
C	4	38.8	27.5	2900	1	US-08-443-015-9	Sequence 9, Appli	
C	5	38.2	27.1	30001	1	US-08-123-468-1	Sequence 1, Appli	
C	6	38.2	27.1	30001	2	US-08-474-933-1	Sequence 1, Appli	
C	7	38	27.0	2310	4	US-09-036-987A-25	Sequence 25, Appli	
C	8	34.2	24.3	43280	2	US-08-804-227C-1	Sequence 1, Appli	
C	9	34	24.1	1485	1	US-07-702-771-2	Sequence 2, Appli	
C	10	34	24.1	1485	1	US-08-438-983-2	Sequence 2, Appli	
C	11	34	24.1	1485	1	US-08-741-1485-2	Sequence 2, Appli	
C	12	34	24.1	2685	2	US-08-935-450-5	Sequence 5, Appli	
C	13	33.8	24.0	1212	2	US-08-443-933-5	Sequence 5, Appli	
C	14	33.8	24.0	8937	2	US-08-443-933-1	Sequence 1, Appli	
C	15	33.8	24.0	10706	1	US-08-411-389-1	Sequence 1, Appli	
C	16	33.4	23.7	2634	3	US-08-911-853-30	Sequence 30, Appli	
C	17	33.4	23.7	17612	3	US-08-911-853-29	Sequence 29, Appli	
C	18	33.2	23.5	1215	3	US-08-911-853-29	Sequence 8, Appli	
C	19	33.2	23.5	17612	3	US-08-911-853-29	Sequence 29, Appli	
C	20	32.8	23.3	1032	1	US-08-400-422-2	Sequence 2, Appli	
C	21	32.8	23.3	1951	3	US-08-923-865-1	Sequence 1, Appli	
C	22	32.6	23.1	862	1	US-08-373-280-4	Sequence 4, Appli	
C	23	31.8	22.6	4649	6	5183745-1	Patent No. 5183745	
C	24	31.8	22.6	6443	6	5183745-5	Patent No. 5183745	
C	25	31.8	22.6	20235	1	US-07-642-734C-3	Sequence 3, Appli	
C	26	31.8	22.6	20235	3	US-08-439-009A-3	Sequence 3, Appli	
C	27	31.6	22.4	900	5	PCR-US95-04801-3	Sequence 3, Appli	

ALIGNMENTS

```

RESULT      1
US-09-009-119-1/c
; Sequence 1, Application US/09009119
; Patent No. 6160206
; GENERAL INFORMATION:
; APPLICANT: SATO, Ryo
; APPLICANT: BOYNTON, John E.
; APPLICANT: GILLHAM, Nicholas W.
; APPLICANT: HARRIS, Elizabeth W.
; TITLE OF INVENTION: Porphyrin Accumulating-Type Herbicide Resistance Gene
; FILE REFERENCE: substitute sequence listing
; CURRENT APPLICATION NUMBER: US/09/009,119
; CURRENT FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-09-009-119-1

```

Query Match	98.9%	Score 139.4	DB 4	Length 3381
Best Local Similarity	99.3%	Pred. No. 7.1e-25		
Matches 140	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	gcgcgcagagccctgggtctcttcgaactacccgcgcgtggcgccgtgaacctgtcgttac	60	
Db	3127	GCGC CGCAGGCCCTGGGGCTCTTCGACTACCCGCCGATGGCGCGGTGACGCTCGCTAC	3068	
QY	61	ccgctcgagccctcgcgagagcgcaagacctcgacgggtcgttcgaggggtcgggtc	120	
Db	3067	CGCTGAGCCCTGCGGAGAGCGCAGGCCCTCGGACGGGTCCGTCGCCGGCTCGGT	3008	
QY	121	cagctgcacccgcgcacgcag	141	
Db	3007	CAGCTGCACCCGCGCACGCAG	2987	

```

RESULT      2
PC1-US95-09098-1/c
; Sequence 1, Application PC/TUS9509098
;
; GENERAL INFORMATION:
;
; APPLICANT: Sato, Ryo
; APPLICANT: Boynton, John E.
; APPLICANT: Gillham, Nicholas W.
; APPLICANT: Harris, Elizabeth H.
; TITLE OF INVENTION: Porphyrin-Accumulating Type Herbicide
; TITLE OF INVENTION: Resistance Gene
; NUMBER OF SEQUENCES: 1
;

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09098
FILING DATE: 20-JUL-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 2185-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3383 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Chlamydomonas reinhardtii
STRAIN: RS-3
PCT-US95-09098-1

Query Match 96.68; Score 136.2; DB 5; Length 3383;
Best Local Similarity 97.98; Pred. No. 3.9e-24;
Matches 138; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 gccgcgagccctggctcttcgactaccgcgcggtggcgccgtgacgtgtcgatc 60
|||||
DB 3128 GCCGCGAGGCCCTGGGCTCTTCGACTACCCGCCGATGGCGCGGTGACGTGCTAC 3069
QY 61 ccgctagccgctgagagagcgcacagccctcgacggttcggtgccggcttcgt 120
|||||
DB 3068 CCCTGAGCGCGCTGGGAGGAGCGACGACCTCGGACGGGTCCGTGCCGGCTTCGGT 3009
QY 121 cagctgaccgcgcacgcag 141
||
DB 3008 CATCTGACCCGCGCAGCAG 2988

RESULT 3
US-08-034-650-9
; Sequence 9, Application US/08034650
; Patent No. 5641671
; GENERAL INFORMATION:
; APPLICANT: BOS, Jannetje W.
; APPLICANT: FRENKEN, Leon G.
; APPLICANT: VERRIJPS, Cornelis T.
; APPLICANT: VISSER, Christiaan
; TITLE OF INVENTION: PRODUCTION OF ACTIVE PSEUDOMONAS GLUMAE
; TITLE OF INVENTION: LIPASE IN HOMOLOGOUS OR HETEROLOGOUS HOSTS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.

COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/034,650
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,235
FILING DATE: 03-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: PNK/5970/91731
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2900 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 483..1556
FEATURE:
NAME/KEY: CDS
LOCATION: 1559..2617
US-08-034-650-9

Query Match 27.58; Score 38.8; DB 1; Length 2900;
Best Local Similarity 55.18; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 2 ccgcgagccctggctcttcgactaccgcgcggtggcgccgtgacgtgtcgatcc 61
|||||
DB 2046 CCGCGGACGCGCGGGGTCTGGCGGACAGCTCGATCGCGCCCATGAGCTCCGCTCG 2105
QY 62 cgcctagccgctgagagagcgcacagccctcgacggttcggtgcggggttcggtc 121
|||||
DB 2106 ATCAGCGCGCGCGCTGGCGGACCGCACGCTCGCGGAGTGGCGCGGCTTCTTCGGCG 2165
QY 122 agctgaccgcgcacgc 139
|||||
DB 2166 ACGAGCAGCGCGGCGAGC 2183

RESULT 4
US-08-449-015-9
; Sequence 9, Application US/08449015
; Patent No. 5804409
; GENERAL INFORMATION:
; APPLICANT: BOS, Jannetje W.
; APPLICANT: FRENKEN, Leon G.
; APPLICANT: VERRIJPS, Cornelis T.
; APPLICANT: VISSER, Christiaan
; TITLE OF INVENTION: PRODUCTION OF ACTIVE PSEUDOMONAS GLUMAE
; TITLE OF INVENTION: LIPASE IN HOMOLOGOUS OR HETEROLOGOUS HOSTS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,015
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,235
FILING DATE: 03-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: PNK/5970/91731
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2900 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 483..1556
FEATURE:
NAME/KEY: CDS
LOCATION: 1559..2617
US-08-449-015-9

Query Match 27.5%; Score 38.8; DB 1; Length 2900;
Best Local Similarity 55.1%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 2 ccgcgagccctggctcttcgaactaccgcgcggtggcgccgtgacgctgtgtacc 61
Db 2046 CCGCGACGCGCGGTGTCGGCGACAAAGTCGATCGCGCGCCGATGCGCTCGCGTCG 2105
QY 62 cgtgtagccgctgctggagggagcgcaaggcctcgacggttcctgcccgggttcggtc 121
Db 2106 ATCAGCGCGCGCGCTGCGCGACCGACGCTCGCGAGTGGCGCGGCGGTCTTCGCG 2165
QY 122 agctgacccgcgcacgc 139
Db 2166 ACGAGCAGCCCGCGCAGC 2183

RESULT 5
US-08-125-468-1
Sequence 1, Application US/08125468
Patent No. 5585385
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Strathy, Nancy
APPLICANT: Fantini, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlortetracycline and tetracycline formation and cosmid
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESS: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,933
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,468

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,468
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)831-3241
TELEFAX: (201)831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-125-468-1
Query Match 27.1%; Score 38.2; DB 1; Length 3001;
Best Local Similarity 54.7%; Pred. No. 0.2;
Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 3 ccgcgagccctggctcttcgaactaccgcgcggtggcgccgtgacgctgtgtacc 62
Db 14962 CGAGTCGCTACGCGCGCGGACCCCTCTGTGCGCGCGGATGCGGACGCCGA 15021
QY 63 gctgagcgcctgctggagggagcgcaaggcctcgacggttcctgcccgggttcggtca 122
Db 15022 GCTGATCATCCGCGGACGCGCGGAGGATCAGCGATTGCTTCGCGCGCGCGGTCT 15081
QY 123 gctgacccgcgcacgcag 141
Db 15082 GCTGCTGCTCCCGCAACAG 15100
RESULT 6
US-08-474-933-1
Sequence 1, Application US/08474933
Patent No. 5866410
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Strathy, Nancy
APPLICANT: Fantini, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlortetracycline and tetracycline
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESS: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,933
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,468

; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,255-02
; TELECOMMUNICATION INFORMATION:
; * TELEPHONE: (201)831-3241
; * TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-474-933-1

Query Match 27.1%; Score 38.2; DB 2; Length 30001;
Best Local Similarity 54.7%; Pred. No. 0.2;
Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 3 cgcagagccctggctcttcgactaccgcgcggtggcgccgtgacgctgtcgtaccc 62
Db 14962 CGAGCTGCGCTACGCGCGCGGCGACACCTCTGTGCGCGCGGATGCGGACGCCGA 15021

Qy 63 gctgagcgcgtgcgggagcgcgaaggcctcgagcgggtccgtgcgggcttcggtca 122
Db 15022 GCTGATCACCGCGGACGCGCGGAGGATGCTTCGCGCGGCGCGCGGTCT 15081

Qy 123 gctgacccgcgcagcag 141
Db 15082 GCTGCTGCTCCCGGACAG 15100

RESULT 7
US-09-036-987A-25
; Sequence 25, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816

; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..1077
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1165..1992
US-09-036-987A-25

Query Match 27.0%; Score 38; DB 4; Length 2310;
Best Local Similarity 57.6%; Pred. No. 0.23;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 1 gccacccagccctggctcttcgactaccgccggtggcgccgtgacgctgtcgtac 60
Db 1213 GTCGCCCGACGCGTGGGTCTCTCCGGAGACGCCGAGACGCGGAGCTGCGACCGGTG 1272

Qy 61 ccgctgagcgcctgcgggagcgcgaaggcctcgagcgggtccggtgccggggttcg 118
Db 1273 GTGATCAACCGCGCGGTACACGCGGTGTGACGCGCGCGGAGTCCGACCGGACCGCG 1330

RESULT 8
US-08-804-227C-1
; Sequence 1, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostek, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 816..14234


```
; NAME: Kanski, Antoinette F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 20344-20553.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141 MRSNFOERS SFO
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1212 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 211..1212
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 52..54
; OTHER INFORMATION: /note= "Upstream in frame stop"
; OTHER INFORMATION: codon
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 98..119
; OTHER INFORMATION: /note= "Oligonucleotide used for
; OTHER INFORMATION: primer extension"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (270-271)
; OTHER INFORMATION: /note= "Position of the first
; OTHER INFORMATION: intron and alternate sequences(SEQ ID NO:6 through SEQ ID NO:
; OTHER INFORMATION: diverge"
; US-08-449-933-5

Query Match 24.0%; Score 33.8; DB 2; Length 1212;
Best Local Similarity 58.4%; Pred. No. 2.2;
Matches 59; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 8 aggcctcggtctcttctactaccgcgcgtggcgccgtgacgtgtcgtagccgtga 67
Db 131 AGCCTCCGCTCCCGCCCTCTTCCGCGCCAGCGCGCCGCCACCCCTCCCTCCGCCGC 190
QY 68 gcgcgtgcggggaggagcaaggcctcgagcgggtccgtg 108
Db 191 CCCCCGCGCGGGGAGGACATGCGCGGCACAGGCGGGTG 231

RESULT 14
US-08-449-933-1
; Sequence 1, Application US/08449933
; Patent No. 5859195
; GENERAL INFORMATION:
; APPLICANT: Collins, Francis S.
; APPLICANT: Wallace, Margaret R.
; APPLICANT: Marchuk, Douglas A.
; APPLICANT: Anderson, Lone B.
; APPLICANT: Guttman, David H.
; TITLE OF INVENTION: Neurofibromatosis Gene
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,933
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kanski, Antoinette F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 20344-20553.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141 MRSNFOERS SFO
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8937 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 17q11.2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6814..8937
; OTHER INFORMATION: /note= "Entire length of sequence
; OTHER INFORMATION: clone p5"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..8646
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8425..8646
; OTHER INFORMATION: /note= "219 nt PstI-HindIII
; OTHER INFORMATION: fragment designated pMAL.B3A"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382..1302
; OTHER INFORMATION: /note= "918 nt HpaI-PstI fragment
; OTHER INFORMATION: designated pMAL.HF3A.P"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382..3909
; OTHER INFORMATION: /note= "3523 nt HpaI-XhoI fragment
; OTHER INFORMATION: designated pMAL.HF3A.X"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8383..8937
; OTHER INFORMATION: /note= "Clone B3A"
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
; TITLE: Type 1 Neurofibromatosis Gene: Correction
; JOURNAL: Science
; VOLUME: 250
; ISSUE: 12/21/90
; PAGES: 1749-
; DATE: 12/21-1990
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 8937
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
; TITLE: Type 1 Neurofibromatosis Gene: Identification
; TITLE: of a Large Transcript in Three NFI Patients
; JOURNAL: Science
; VOLUME: 249
; ISSUE: 07/13/90
; PAGES: 181-186
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2001, 00:00:32 ; Search time 1116.81 Seconds
(without alignments)
1861.953 Million cell updates/sec

Title: US-09-331-723A-4

Perfect score: 141

Sequence: 1 gccgcgagccctgggtc.....agctgcacccgcgcacgcag 141

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_cm:*

8: gb_ov:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_pil:*

13: gb_pil2:*

14: gb_pil3:*

15: gb_pil4:*

16: em_bal:*

17: em_ba2:*

18: em_fun:*

19: em_htgo_hum:*

20: em_htgo_inv:*

21: em_htgo_rnd:*

22: em_htg_hum1:*

23: em_htg_hum2:*

24: em_htg_hum3:*

25: em_htg_hum4:*

26: em_htg_hum5:*

27: em_htg_hum6:*

28: em_htg_hum7:*

29: em_htg_hum8:*

30: em_htg_inv1:*

31: em_htg_inv2:*

32: em_htg_inv3:*

33: em_htg_rnd:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_cm:*

43: em_or:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_pl:*

48: em_ro:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_vi:*

53: gb_stsl:*

54: gb_sts2:*

55: gb_sts3:*

56: gb_sy:*

57: gb_un:*

58: gb_vil:*

59: gb_vil2:*

60: gb_vil3:*

61: gb_vil4:*

62: gb_vil5:*

63: gb_vil6:*

64: gb_vil7:*

65: gb_vil8:*

66: gb_vil9:*

67: gb_vil10:*

68: gb_vil11:*

69: gb_vil12:*

70: gb_vil13:*

71: gb_vil14:*

72: gb_vil15:*

73: gb_vil16:*

74: gb_vil17:*

75: gb_vil18:*

76: gb_vil19:*

77: gb_vil20:*

78: gb_vil21:*

79: gb_vil22:*

80: gb_vil23:*

81: gb_vil24:*

82: gb_vil25:*

83: gb_vil26:*

84: gb_vil27:*

85: gb_vil28:*

86: gb_vil29:*

87: gb_vil30:*

88: gb_vil31:*

89: gb_vil32:*

90: gb_vil33:*

91: gb_vil34:*

92: gb_vil35:*

93: gb_vil36:*

94: gb_vil37:*

95: gb_vil38:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141	100.0	141	9	A78950
2	141	100.0	2480	12	AF068635
3	139.4	98.9	2573	9	A78956
4	139.4	98.9	3381	9	A78965
5	139.4	98.9	10265	12	AF030179
c 6	136.2	96.6	3383	9	A78933
c 7	39.4	27.9	37750	3	SC6D7
c 8	39.4	27.9	151089	13	AP003105
c 9	39	27.7	4767	1	AB021222
10	39	27.7	11403	1	AE004488
11	39	27.7	16903	1	AE004724

```

12 38.8 27.5 2900 3 PGLIP
13 38.8 27.5 2900 9 A16323
14 38.8 27.5 2900 9 AR038049
15 38.8 27.5 2900 10 I49947
16 38.4 27.2 7459 3 SCBA2
17 38.2 27.1 4422 3 STWCTS
18 38.2 27.1 30001 9 AR031172
19 38.2 27.1 30001 10 I32939
20 37 26.2 4674 2 AL047026
21 37 26.2 146921 13 AP002836
22 37 26.2 158133 13 AP002816
23 36.8 26.1 3777 1 AF139661
24 36.8 26.1 40744 3 SCJ4
25 36.6 26.0 9413 2 LDMGFA
26 36.6 26.0 10506 1 AE004806
27 36.6 26.0 11979 3 MXCRTNODA
28 36.6 26.0 40476 3 SCC75A
29 36.6 26.0 63554 14 H0410G08
30 36.4 25.8 12172 1 AE001872
31 36.4 25.8 13431 1 AE005044
32 36.4 25.8 33517 3 SC10B7
33 36.4 25.8 34158 3 SC2K31
34 36.4 25.8 40007 3 SC4B10
35 36.4 25.8 43632 3 SC6A5
36 36.4 25.8 281026 76 AC079572
37 36.2 25.7 938 6 LMB19KP
38 36.2 25.7 32941 3 SCM10
39 36 25.5 1026 9 A70058
40 36 25.5 3217 3 SGSECA
41 36 25.5 6231 3 SCC117
42 36 25.5 8168 2 D87924
43 36 25.5 12198 1 AE001953
44 36 25.5 26500 3 SC9H11
45 36 25.5 36224 3 SCD78

```

ALIGNMENTS

```

RESULT 1
A78950
LOCUS
DEFINITION Sequence 4 from Patent WO9829554.
ACCESSION A78950
VERSION A78950.1 GI:6092078
KEYWORDS
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
AUTHORS Ishige,F. and Sato,R.
TITLE METHODS OF CONFERRING PPO-INHIBITING HERBICIDE RESISTANCE TO PLANTS
BY GENE MANIPULATION
JOURNAL Patent: WO 9829554-A 4 09-JUL-1998;
ISHIGE FUMIHARU (JP); SATO RYO (JP)
FEATURES
    source
        1..141
            /organism="Chlamydomonas reinhardtii"
            /strain="CC-407"
            /db_xref="taxon:3055"
            15 a 53 c 54 g 19 t
BASE COUNT 15 a 53 c 54 g 19 t
ORIGIN

```

```

Query Match 100.0%; Score 141; DB 9; Length 141;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggcgcgagccctggctccttcgactaccgcgcggtggcgccgtgacgtgtcgatc 60
Db 1 ggcgcgagccctggctccttcgactaccgcgcggtggcgccgtgacgtgtcgatc 60

```

```

QY 61 ccgctgagccctggcgagagcgcaagcctcgagcggtccgctccgctcggttcgggt 120
Db 61 CCGCTGAGCCCTGGCGGAGGAGCGCAAGCGCTCGGACGGTCCGCTCCGGGCTTCGGT 120
QY 121 cagctgcacccgcgcacgcag 141
Db 121 CAGCTGACCCCGGCACGAG 141

```

RESULT 2

```

AF068635 2480 bp mRNA PLN 27-NOV-1998
LOCUS Chlamydomonas reinhardtii protoporphyrinogen oxidase precursor
DEFINITION (Ppx1) mRNA, nuclear gene encoding chloroplast protein, complete.
CDs
ACCESSION AF068635
VERSION AF068635.1 GI:3928793
KEYWORDS
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
AUTHORS Randolph-Anderson,B.L., Sato,R., Johnson,A.M., Harris,E.H.,
Hauser,C.R., Oeda,K., Ishige,F., Nishio,S., Gillham,N.W. and
Boydton,J.E.
TITLE Isolation and characterization of a mutant protoporphyrinogen
oxidase gene from Chlamydomonas reinhardtii conferring resistance
to porphyrin herbicides
JOURNAL Plant Mol. Biol. 38 (5), 839-859 (1998)
MEDLINE 99077317
REFERENCE 2 (bases 1 to 2480)
AUTHORS Randolph-Anderson,B.L., Sato,R., Johnson,A.M., Harris,E.H.,
Hauser,C.R., Oeda,K., Ishige,F., Nishio S., Gillham,N.W. and
Boydton,J.B.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1998) DCMB Group, LSRC Building, Duke University,
Research Drive, Durham, NC 27708, USA
FEATURES
    source
        1..2480
            /organism="Chlamydomonas reinhardtii"
            /strain="CC-621"
            /db_xref="taxon:3055"
            /map="13.7 map units from srl, 12.3 map units from pfl6"
            /clone_lib="Lambda Zap II library; Wafenschmidt et al
            (1993) Plant Cell 5:809-820"
            /chromosome="linkage group IX"
            /note="wildtype strain"
            1..2480
            /gene="Ppx1"
            91..1782
            /gene="Ppx1"
            /function="prophyrin biosynthesis"
            /codon_start=1
            /product="protoporphyrinogen oxidase precursor"
            /protein_id="AAC79685.1"
            /db_xref="GI:3928794"
            /translation="MLTQPTGATSSRRSIOIRSAHVSAKVAPRPTFVSAPATA
            ASPATAARRTLHETAAANTCAPTASGACVAKTLDNVYDVIIVGGSLGVTQALAA
            QKIONFLVTEAREVGNITSGSGYVWEEGPNFQPNDSMLQIADVSGCEKDLVF
            GPTAFRVEGKLRVPVPSGLDAFTDLSIPKIRAGLGAIGLINGAMPSEESVE
            QTRNLGDEVFFRLIEFFCSGVYAGDPSKLSKAAFRNIWLEKNGGSLVGAIKLF
            QERQSNAPPDRPLPKPKGTQVSGFRGLKMLPDATERNIPDKIRVNNKVLGKRE
            ADGRYGLVDTPEGRVKVFAVALTAPSVVADLVKEQAPAAALGSDYDPVGVG
            TLYPLSAVREERKASDGSVPGFQGLHPTQTGTTLTITYSLSFFPGRAPEGHMLLN
            YIGGTTNRGINOTTLEQVQDKLRNMVVKPDAPKPRVGVWRVRAIPQPNLGHLL
            EQLDKARKALDAAGLQGVHLGGNVSVSGALGVVEHGYESAANLAKSVSKAAVKA"
BASE COUNT 462 a 758 c 834 g 426 t
ORIGIN

```

```

Query Match 100.0%; Score 141; DB 12; Length 2480;
Best Local Similarity 100.0%; Pred. No. 4.6e-15;

```

87 cagcgcacccgcgcgcag 141
 |||||
 Db 1339 CAGGTCACCCGCGCACGCAG 1359

ORGANISM *Chlamydomonas reinhardtii*
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; *Chlamydomonas*.

```

mrna
join<1717. 2163,2366. 2410,2983. 3124,3393. 3442,
3800. 3885,4153. 4221,4391. 4449,4801. 4870,5159. 5312,
5709. 5855,6113. 6256,7241. 7333,7622. 7718,8103. >8191)
/genes="Ppx1"
/product="herbicide-resistant protoporphyrinogen oxidase
precursor"

```

gene
CDS
*
transit_peptide
mat_peptide
variation
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 140; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Y 1 gccgcgagccctggctcttcctgactaccgcggtggtggcgccgtgacgtctcgtac 60
Db 5715 GCGCGGAGCCCTGGCTCTTCGACTACCGCGCGATGGCGCGTGCCTGCTAC 5774
Y 61 ccgctgagccgtgaggagagcgaagcctcgagcgtcggtgacgtctcgtac 120
Db 5775 CCCTGAGCGCGTGGGAGGAGCGCAGGCGCTCGGACGGTCCGTGCGGCTTCGT 5834
Y 121 cagctgaccgcgcgcagcag 141
Db 5835 CAGCTGACCGCGCGACGAG 5855
RESULT 6
LOCUS A79833/c 3383 bp DNA PAT 20-OCT-1999
DEFINITION Sequence 1 from Patent WO9704088.
ACCESSION A79833
VERSION A79833.1 GI:6092737
KEYWORDS Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
REFERENCE Sato,R. and Boynton,J.
AUTHORS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
TITLE Chlamydomonas reinhardtii.
JOURNAL Chlamydomonas reinhardtii.
FEATURES
source 1.3383
/organism="Chlamydomonas reinhardtii"

<1717..>8191
/gene="Ppx1"
join(1717..2163,2366..2410,2983..3124,3393..3442,
3800..3885,4153..4221,4391..4449,4801..4870,5159..5312,
5709..5855,6113..6256,7241..7333,7622..7718,8103..8191)
/gene="Ppx1"
/function="porphyrin biosynthesis"
/codon_start=1
/product="herbicide-resistant protoporphyrinogen oxidase
precursor"
/protein_id="AAC79630.1"
/db_xref="GI:3928569"
/translation="MWLTQPTCTATASSRBSQIRSAHVSAKVAPRPTPFVSPATA
ASPTAAARTLHRTAAATGAPTASGAGYAKTLDNVYDVLVGGVLSGLVGCALAA
QKIQNLFTEARVGRVGGTGVWEEGPNFQNDMSLQIADVSCCKDLVF
GDTPTAFVWEGKRLPVPFSGDFTDLMSIFPKIRAGLGAIGLINGAMSPFEESVE
QFTRNLGDFVFRLEIPFCGVSAGVAGPDKSLMKAAFNWILEKNGSLVGGAKLIF
QERQSNAPRPPRLPPKPGQVTGVSFRKGLKMLPDAIERNIPDKIRVNWKLVSIGRE
ADGRYGLVDTPEGRVFAVAVALTAPSVVADLVKEQAPAAAEALFDFYPMGAV
TLSPVLSAVREERKASDGVFGQLHPRVQITTLGTTIYSSSLPFGRAPEGHMLIN
YIGSTNRGIVNQTTEOLVEQVDKDLNRMVVKPDAPKRVGVVRVPRAPVPOFNIGHL
EQDKARKALDAAGLQVHLGSLGVNYSVALGVKVEHGESANLAKSVSKAAVKA"
/transit_peptide 1717..1962
/gene="Ppx1"
join(1963..2163,2366..2410,2983..3124,3393..3442,
3800..3885,4153..4221,4391..4449,4801..4870,5159..5312,
5709..5855,6113..6256,7241..7333,7622..7718,8103..8188)
/gene="Ppx1"
/product="herbicide-resistant protoporphyrinogen oxidase"
5751
/gene="Ppx1"
/note="encodes methionine rather than wildtype valine"
/replace="g"
BASE COUNT 1903 a 3057 c 3413 g 1892 t
ORIGIN

Query Match 98.9%; Score 139.4; DB 12; Length 10265;
Best Local Similarity 99.3%; Pred. No. 4e-15;
Matches 140; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Y 1 gccgcgagccctggctcttcctgactaccgcggtggtggcgccgtgacgtctcgtac 60
Db 5715 GCGCGGAGCCCTGGCTCTTCGACTACCGCGCGATGGCGCGTGCCTGCTAC 5774
Y 61 ccgctgagccgtgaggagagcgaagcctcgagcgtcggtgacgtctcgtac 120
Db 5775 CCCTGAGCGCGTGGGAGGAGCGCAGGCGCTCGGACGGTCCGTGCGGCTTCGT 5834
Y 121 cagctgaccgcgcgcagcag 141
Db 5835 CAGCTGACCGCGCGACGAG 5855
RESULT 7
LOCUS SC6D7/c 37750 bp DNA BCT 02-MAY-2000
DEFINITION Streptomyces coelicolor cosmid 6D7.
ACCESSION AL133213
VERSION AL133213.2 GI:7688318
KEYWORDS acd; acyl-CoA dehydrogenase (EC 1.3.99.-); ankyrin-like protein;
asnC-family transcriptional regulator; ATP
phosphoribosyltransferase; cbxX/CifX family protein; chdI;
chitinase; glycosyl transferase; GntR-family regulatory protein;
hlsG; integral membrane protein; membrane protein;
phosphodiesterase; secreted protein; tetr-family transcriptional
regulator; transmembrane transport protein.
SOURCE Streptomyces coelicolor A3(2).
ORGANISM Streptomyces coelicolor A3(2).
REFERENCE Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
AUTHORS Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
TITLE 1 (bases 1 to 37750)
JOURNAL Redenbach,M., Kieser,H.M., Denapalte,D., Eichner,A., Cullum,J.,
MEDLINE Kinashi,H. and Hopwood,D.A.
AUTHORS A set of ordered cosmids and a detailed genetic and physical map
JOURNAL for the 8 Mb Streptomyces coelicolor A3(2) chromosome
REFERENCE Mol. Microbiol. 21 (1), 77-96 (1996)
JOURNAL 97000351
AUTHORS 2 (bases 1 to 37750)
JOURNAL Seeger,K.J. and Harris,D.
REFERENCE 3 (bases 1 to 37750)
JOURNAL Thomson,N.R., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
AUTHORS Direct Submission
TITLE Submitted (24-NOV-1999) Streptomyces coelicolor sequencing project,
JOURNAL Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood. [3] John Innes Centre, Norwich Research Park,
Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT On May 3, 2000 this sequence version replaced gi:6468696.
Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded
by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are
available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S.coelicolor/)
CDS are numbered using the following system eg SC7B7.01c, SC (S.
coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.

/strain="RS-3"
/db_xref="taxon:3055"
BASE COUNT 671 a 1008 c 1057 g 647 t
ORIGIN
Query Match 96.6%; Score 136.2; DB 9; Length 3383;
Best Local Similarity 97.9%; Pred. No. 2.4e-14;
Matches 138; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Y 1 gccgcgagccctggctcttcctgactaccgcggtggtggcgccgtgacgtctcgtac 60
Db 3128 GCGCGGAGCCCTGGCTCTTCGACTACCGCGCGATGGCGCGTGCCTGCTAC 3069
Y 61 ccgctgagccgtgaggagagcgaagcctcgagcgtcggtgacgtctcgtac 120
Db 3068 CCGCTAGCCCGCTGGGAGGAGCGGACGACCTCGGACGGTCCGTGCGGCGGCTTCGGT 3009
Y 121 cagctgaccgcgcgcagcag 141
Db 3008 CATCTGACCGCGCGACGAG 2988
RESULT 7
LOCUS SC6D7/c 37750 bp DNA BCT 02-MAY-2000
DEFINITION Streptomyces coelicolor cosmid 6D7.
ACCESSION AL133213
VERSION AL133213.2 GI:7688318
KEYWORDS acd; acyl-CoA dehydrogenase (EC 1.3.99.-); ankyrin-like protein;
asnC-family transcriptional regulator; ATP
phosphoribosyltransferase; cbxX/CifX family protein; chdI;
chitinase; glycosyl transferase; GntR-family regulatory protein;
hlsG; integral membrane protein; membrane protein;
phosphodiesterase; secreted protein; tetr-family transcriptional
regulator; transmembrane transport protein.
SOURCE Streptomyces coelicolor A3(2).
ORGANISM Streptomyces coelicolor A3(2).
REFERENCE Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
AUTHORS Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
TITLE 1 (bases 1 to 37750)
JOURNAL Redenbach,M., Kieser,H.M., Denapalte,D., Eichner,A., Cullum,J.,
MEDLINE Kinashi,H. and Hopwood,D.A.
AUTHORS A set of ordered cosmids and a detailed genetic and physical map
JOURNAL for the 8 Mb Streptomyces coelicolor A3(2) chromosome
REFERENCE Mol. Microbiol. 21 (1), 77-96 (1996)
JOURNAL 97000351
AUTHORS 2 (bases 1 to 37750)
JOURNAL Seeger,K.J. and Harris,D.
REFERENCE 3 (bases 1 to 37750)
JOURNAL Thomson,N.R., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
AUTHORS Direct Submission
TITLE Submitted (24-NOV-1999) Streptomyces coelicolor sequencing project,
JOURNAL Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood. [3] John Innes Centre, Norwich Research Park,
Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT On May 3, 2000 this sequence version replaced gi:6468696.
Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded
by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are
available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S.coelicolor/)
CDS are numbered using the following system eg SC7B7.01c, SC (S.
coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.

The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.go.jp/junc/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or attg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: this sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.

Cosmid 607. Overlaps with cosmid 1A8A on the AseI-g genomic restriction fragment.

FEATURES

source
1. .37750
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"
/clone="cosmid 6D7"
complement(1..4276)
/gene="SC6D7.32"
complement(<1..4276)
/gene="SC6D7.32"
/note="SC6D7.32, unknown, len: > 1425 aa. Weakly similar to several including: Streptomyces coelicolor
TR:Q95A4(EMBL:AL035478) putative secreted protein
SC265.19 (1156 aa), fasta scores opt: 327 z-score: 314.0
E(): 4.1e-10 28.0% identity in 938 aa overlap and
Rhodococcus capsulatus TR:068032(EMBL:AF010496) potential
exonuclease (1238 aa), fasta scores opt: 439 z-score:
421.6 E(): 4.1e-16 28.8% identity in 1154 aa overlap. The
predicted protein product of this CDS is alanine rich and
contains multiple probable coiled-coils between residues:
369 to 401 (33 residues), 406 to 433 (28 residues), 483 to
519 (37 residues), 563 to 604 (42 residues), 850 to 921 (72
residues), 926 to 982 (57 residues), 1025 to 1054 (30
residues), 1072 to 1106 (35 residues), 1124 to 1167 (44
residues), 1199 to 1232 (34 residues) and 1278 to 1309 (32
residues)."
/codon_start=1
/transl_table=11
/product="hypothetical protein SC6D7.32."
/protein_id="CAB61684.1"
/db_xref="GI:6468728"
/translation="MYELSRVRLYSIGPAGARYADTVLDLGVGRVLPDPAPTAQEEFF
EEFVGPVRRPAPAGVLFLENGGKSVLLKLIIFSVMLPGHRTNLGGSSGVLRFLFLA
DCGHVALEWHQVOTCECVVYKYSWGRQVSDPKFAEAWYSFRPGGLTLDNLP
VAETAAPVPEGASCAQGRERTMKGFDPALTAGKAYPHLEWHWEIHDWJELHGD
LGLDPELFYQREMANDEGEANGLFVYKSDFTDLLRAVTDTRDTGLADLVGGCG
NKGRRAGLIAERDFTAGSDLLGRTVEAEARTRADVHTAEARTRNLARLTARG
AQEVRAGDLAQKRVTAAYAVYTHAESANDRNLIAELAYRHASLTLGABKAARQK
RELAERTLSAWQAAEAVLRHRAADRVARVSAIQEAPERDAAPALAAARAAVALV
RTLHAAQKAHEHNEEARSAAQAEAGDAHRDSTTAETAQARSAGHLRQRLTE
VEQTAFAVAGVLDDSPADPAALAAASDAEKTVAADWTALLETAGRATHEAREA
AAESRELTAAARAADAAPRAHEARTATAELAEERLAEGLAEARPGIPOPRA
RDADPRAARTAADGALSPELDFADELRELDLDDTVSSAEROLFOLRTAAADDSR
ILGLIITDPSGHTPAREVIGDAAALPRSAVGTAAALLAPAPAGSGDGVLPFPNP
AMHDEQAADERHALRAGERDEEIRALARIQKQLELAARLNSWRTGCPAGRLTEL
ARTAE
LAFLRERAGQVRLRELADEGAEEAQAQCLERAADDERAAQRAADDAARTRAR
ALRARESEIAGDPDVEAPAGGAEAPKAPLDLREAVRAASQVYKVGAGDARAE
MASSAEOLGRCEAEHRAPEGDTHDLPPELLPDAEHAQSLRTAEVLTITTEA
LARSREHAELLDAHRAEDAAGGDETAAMLRLDLRDHAPAEQDEPEPAGDLEA
RRAAAEARRSLRGCAADLSAAEAVREASOVLVRHANSRTYEDVTPAPQOIRELPAP
ALPFAAKWADAFAPRLVLTDELAUQERNDKSDIVDLRLGLVESALATLRSARLSRL

gene
complement(4382..5263)
/gene="SC6D7.31"
CDS
complement(4382..5263)
/gene="SC6D7.31"
/note="SC6D7.31, possible membrane protein, len: 293 aa.
Contains a possible membrane spanning hydrophobic
domain."
/codon_start=1
/transl_table=11
/product="putative membrane protein."
/protein_id="CAB61683.1"
/db_xref="GI:6468727"
/translation="MTEYVDRESEPEAGAPATAVTPADARADARLVAFLGLOPKLLPA
RQDEYADLLRRYRDPFFARLADAVAGGLIVLEYSPRAGMAVSANESVAVRWGD
YARRAATDGRDFLHGLAHVAAMAYPRPEDLADGGYIGRVTVNGVDFVAFVQACRRLL
ERAPQOGENTDPAUHLAAWAIWRWRSATGATKARLPGSTTIGVKAAGFLT
ESGFLQRTGDDHGGTYTETARYQLQVRDMAGSAALTELLDGLGVVPTDGTATLLPAEE
TDDLELVADAGLPFHS"
complement(5272..5276)
RBS
complement(5352..6878)
gene
/gene="SC6D7.30"
CDS
complement(5352..6878)
/gene="SC6D7.30"
/note="SC6D7.30, possible membrane protein, len: 508 aa.
Contains a possible membrane spanning hydrophobic domain
and a probable coiled-coil between residues 259..286 (28
residues)."
/codon_start=1
/transl_table=11
/product="putative membrane protein."
/protein_id="CAB61682.1"
/db_xref="GI:6468726"
/translation="MTSPASTSDVSPQLSAGPRPAPGPADEGLARLRALACAPLH
DDDAKANKLAGEYSYGMAEVALAALDLVTLNMDPDTGADHDQIVARLIPRVAQAAPH
RYAARHVAWNLVGLNINVSVDGRFVAVYFGDGTGYVRDYDFKLEEVPGQGV
TVLRTTEAVNLVGLADTDTVSAOIAAEVKLEVLINRGLADAAQAAEQARYTVQ
YSETLRRALDTRNRVAVDLSAVPDMLAEALDHADVDRHENAIIITNRKARDTE
DEHKRRAAELVDIVKCIIRRHQQLQSLRLEAGPLFRAEQDQAFATPLTSLGLDLXG
HLVAVPLPVESSALRVTDFAFGTGLKRTFVSVRVGDLVILITPPVERHLAGEMP
EPDLTDPDPSRFESEQLAAAMELLDHPDAPRUSGLLADARADALPYLVALLAA
HAASPVPVATYRQGEKPLFAVDGDTDELDDPEFGGADLVITALLDAAGMAADRTAA
"
complement(6987..7682)
gene
/gene="SC6D7.29"
CDS
complement(6987..7682)
/gene="SC6D7.29"
/note="SC6D7.29, putative GntR-family transcriptional
regulator, len: 231 aa. Similar to many other regulators
e.g. Bacillus licheniformis SW:GNTB_BACLI(EMBL:D31631)
glucanase operon transcriptional repressor (P28 protein)
(243 aa), fasta scores opt: 183 z-score: 234.8 E():
1.1e-05 31.2% identity in 138 aa overlap and Chelatobacter
heintzii SW:NTRA_CHEHE(EMBL:D39411) NTA operon
transcriptional regulator (210 aa), fasta scores opt: 222
z-score: 284.0 E(): 1.9e-08 31.8% identity in 195 aa
overlap. Contains a PS00043 Bacterial regulatory proteins,
gntR family signature and a Pflam match to entry PF00392
gntR, Bacterial regulatory proteins, gntR family. Contains
a putative helix-turn-helix motif between residues 44..65
(+2.58 SD)."
/codon_start=1
/transl_table=11
/product="putative GntR-family transcriptional regulator."
/protein_id="CAB61681.1"
/db_xref="GI:6468725"
/translation="MAEQLTGLADRRALLGRTSTAERVSILRSRIAGEYFPFPTRLS
EDISGIGLGGPPLDGLAAEAVDGRSGREGDMKVGTAIHFHRELVALAGSRTEDEL
AVVRGIGPPYPLDGLAAEAVDGRSGREGDMKVGTAIHFHRELVALAGSRTEDEL
MRSVFAELRLAHFHVVDPRRLHPEYIARNAILSTLESGEREAARLAVYLENSLKR
VVEVILRRVGGEG"
complement(7437..7613)
misc_feature

CDS	1850..3097 /gene="orf2" /note="putative" /codon_start=1 /transl_table=11 /product="chain length factor" /protein_id="BAA87908.1" /db_xref="GI:6518513"	REFERENCE AUTHORS	1 (bases 1 to 11403) Stover, C.K., Pham, X.Q., Erwin, A.L., Miziochuchi, S.D., Warriner, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltzy, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., and Paulsen, I.T. Complete genome sequence of <i>Pseudomonas aeruginosa</i> PA01, an opportunistic pathogen Nature 406 (6759), 959-964 (2000) 2043737
gene	3153..3419 /gene="orf3" 3153..3419 /gene="orf3" /note="putative" /codon_start=1 /transl_table=11 /product="acyl carrier protein" /protein_id="BAA87909.1" /db_xref="GI:6518513"	REFERENCE AUTHORS	2 (bases 1 to 11403) Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Miziochuchi, S.D., Warriner, P., Hickey, M.J., Brinkman, F.S.D., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltzy, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V. Direct Submission Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA Location/Qualifiers 1..11403 /organism="Pseudomonas aeruginosa" /strain="PA01" /db_xref="taxon:287" 168..1946 /gene="PA0508" 168..1946 /gene="PA0508" /codon_start=1 /transl_table=11 /product="probable acyl-CoA dehydrogenase" /protein_id="AAG03897.1" /db_xref="GI:9946373"
gene	3460..4416 /gene="orf4" 3460..4416 /gene="orf4" /note="putative" /codon_start=1 /transl_table=11 /product="cyclase" /protein_id="BAA87910.1" /db_xref="GI:6518514"	FEATURES source	168..1946 /gene="PA0508" 168..1946 /gene="PA0508" /codon_start=1 /transl_table=11 /product="probable acyl-CoA dehydrogenase" /protein_id="AAG03897.1" /db_xref="GI:9946373" translation="MADYKAPLRDMRFLNVEFVSRLWAOLPALAEVDAETAAIL EEAGKTACTIAPLNRECGDEGONNAGVSPAGPEAYRTAEGWVGVDGAYG GMCKVISAQVEELVNSANLFGILPMLFAGACIALAHASDELKDKYLPNMYAGIW AGSMCTEPEDLDGIIIRTAEPQADGSKTIGTKITIGTGHDLTENTIIHLVLAKL PDAPAGKGISILFVPLVKNADGSLGKNSGCSIEHKMGIKASATCYMNDGATG WLYGKIVLAAMFTMMNYRLGVLGTLGATGERSYQSAIYARERIOSRAPTPGVAK DKAADP11VHPDVRMLTLMKALNEGSRFSSVAMQDLTAKYSEDATVTKRAEELVA LLPVPKAPLTDMGLTTHGQITFGHGFIEMWGEQILFIREWQEDCRITQIYETNGIQA DLVRKVIIGSGAFSRHFTDEIKAFVADADEALGFESKPLAAAVENLELTAWLLDRA KGNPNEIGASVVEYLHVFGYTAYMMALMARTALAKAQEDDFVASKLGTAIFYFARL LPRIHSLSASVPAGSESLYLLDAEQF" complement(2277..3758) /gene="nirN" /note="PA0509" complement(2277..3758) /gene="nirN" /codon_start=1 /transl_table=11 /product="probable c-type cytochrome" /protein_id="AAG03898.1" /db_xref="GI:9946374"
CDS	568 a 1877 c 1754 g 568 t BASE COUNT ORIGIN	CDS	complement(2277..3758) /gene="nirN" /note="PA0509" complement(2277..3758) /gene="nirN" /codon_start=1 /transl_table=11 /product="probable c-type cytochrome" /protein_id="AAG03898.1" /db_xref="GI:9946374" translation="MRLIGLALGILLCALAQAGEAPGALYRHCQACHAGRLGGSG PTLLPESLRKPAAQREVILHGPATPMAGFAGQDDAADALVALVIAQAPPEQW SAEDIRASQVPHPLATPSRPEADPLNLFVVEGSDHVTILDGDFREPAREPQS RYALHGGKFSFDPGRVYFASRGDWTLVDLYNLKVAEVRAGLNTLAVSODGRW LVGNLPGNLVLLDARDLSLVQIPADAGQASRVSAVYTPAPRHSFVVALKDVHEL WELPYANGKPVAKRLVADYLDLDFSPDPYRYLLGSSRAGGGEVTELDGSGARVASI PLSGMPLGSGTYKKRGRWVFATPNISRGVTSVIDLONWKPKLEIVTDGPGFMESH ADSPYATDTFLGKKHDEILLIDKQTLTAHLRPSGKVGAGHVEFTRDGRYALLSW DRGALVYVDAHSLEEVKRLPMKPSGKYNVGKNGIYAEGTSH" complement(3749..4588) /gene="PA0510" complement(3749..4588) /gene="PA0510" /codon_start=1 /transl_table=11 /product="probable uroporphyrin-III c-methyltransferase"
Query Match	27.7%; Score 39; DB 1; Length 4767;	gene	
Best Local Similarity	59.5%; Pred. No. 2.5e+02;	CDS	
Matches	66; Conservative 0; Mismatches 45; Indels 0; Gaps 0;		
QY	3 cgcgagccctggctcttcgactaccgcgcggtggcgcggtgacgctgtgaccc 62		
Db	75 CGAAGCGGCTGCCCGCCGCTGCTGTTGGTGGCGGCTGCCCGCGCGGACGC 134		
QY	63 cctagcgcctggagagagcgaagcctcgacgagcgtcgctgcggg 113		
Db	135 GCGTCGACCTGCCGACGCCCGCGGCGGCGGCGGACGAGACCTCTCGC 185		
RESULT 10			
AE004488			
LOCUS	AE004488 11403 bp DNA BCT 30-AUG-2000		
DEFINITION	Pseudomonas aeruginosa PA01, section 49 of 529 of the complete genome.		
ACCESSION	AE004488 AE004091		
VERSION	AE004488.1 GI:9946372		
KEYWORDS	Pseudomonas aeruginosa.		
SOURCE	Pseudomonas aeruginosa		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.		

```
/protein_id="AAG03899.1"
/db_xref="GI:9946375"
/translation="MNTTIPPSLLDVFDPAGSVALVAGPGDPGLLTLRWALLQQA
EVVYVGLARELIALPESQRIYVGRKCGHSLPQEEINELLVRLARQRRVRLK
GGVDFIFGRAGELERLLEAGVQCVPGVTAASGCTVAGIPLTHRLDAQSCITFTVG
HLQNDGRLLDNLWAGLARKQTLVFYMGILGNLAETIARLVEHGLASDTPAALVSGQTQA
AQVTRGALAEPLALARYQLKPTLLVQVVALFAERAMAHPSYLGAGSPYSREAV
ACA"
complement(4599..5762)
/gene="nirJ"
/feature="nirJ"
/feature_note="PA0511"
complement(4599..5762)
/gene="nirJ"
/codon_start=1
/transl_table=11
/product="heme d1 biosynthesis protein NirJ"
/protein_id="AAG03900.1"
/db_xref="GI:9946376"
/translation="MLRISHYLSRLTEDAPTPRRAGRRAPVVIWNLRRCNLTCKHC
YSTSADSPFRGELETAEILRIDDLAGRVRLTILSGGEPLMHPDLFETIAAHARQAGM
FVALSSNGTILDEGNIQVAEAREDFYVIGLIDGRETHDRFRQKQSGFDAAALAMRLC
READIRVGMRTITLTENAAQLPALIDLMRELDVQKFEYLSHLNYSGRRRSRALDAHRH
RTREALALLFERADODIRGSDSDFTVGNNDADAILLDLKRRRPPQOLARLLELLID
WGNASGEGTANIDTGVHPDSYWHHSVGNIRHORFADFWRFPDPLLLQLRPRR
PVVGRSCQHWLDCNGNTRANAGGELWGEDPDCYLSQDEIGLIERIALHAV"
complement(5756..6271)
/gene="PA0512"
complement(5756..6271)
/gene="PA0512"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG03901.1"
/db_xref="GI:9946377"
/translation="MSACISPSDALLARRLIETQAGLELVADPWANATAQLRSEAE
TLALLKLRDAGVIRIRAAVPHNRNGTVMVDVADERIQLRVLGGLSFYSH
CYRRPLRPPWRYNLFAVHGSEAEIERYRQRIILLGDCRADEMLVSSRIILKTTG
LRLAQKEERC"
complement(6246..6689)
/gene="PA0513"
complement(6246..6689)
/gene="PA0513"
/codon_start=1
/transl_table=11
/product="probable transcriptional regulator"
/protein_id="AAG03902.1"
/db_xref="GI:9946378"
/translation="MDEFDRLLNRLOHGLPLRHPYVALLAAELDCREEDILRLDOL
LDDCTLFRGPLFDIPLGGAFTLAASVPEARFEETALLAGWPQVAHNYRREHALN
MWLVVACDSPAEVAETLARLERESGLAVLDLPKREYHYHVLHPEPL"
complement(6682..7206)
/gene="nirL"
/feature="nirL"
/feature_note="PA0514"
complement(6682..7206)
/gene="nirL"
/codon_start=1
/transl_table=11
/product="heme d1 biosynthesis protein NirL"
/protein_id="AAG03903.1"
/db_xref="GI:9946379"
/translation="MNPVEPLAAPORQHLRYLLEQGLPLASPRYRLAERTGAGEDEV
LEQVRNDEGLFRFEGVILHHRALGYTANMLVDVDAEADVAGRALAHETIVSLC
YRRPRRLPMWPNYLFMCIHGRGERGEVERQIETALLERHALRQTHRWLFLSRAYKQCGG
RYTAPPADLERRHG"
complement(7203..7655)
/gene="PA0515"
complement(7203..7655)
/gene="PA0515"
/codon_start=1
/transl_table=11
/product="probable transcriptional regulator"
/protein_id="AAG03904.1"
/db_xref="GI:9946380"
```

```
/translation="MDDLRRLLARYQKGLPICAEPYRRMAETLGCSEAEVLERLRL
EADGLSRVGPVLRHORAGASTLAALAYPERLQVRAERISQYAEVNNYHQREHYNL
WFVITAGDRAQLDRVLAEADTGLQPLDLPMQEAICYDIAFFLEASR"
complement(7664..8842)
/gene="nirF"
/feature="nirF"
/feature_note="PA0516"
complement(7664..8842)
/gene="nirF"
/codon_start=1
/transl_table=11
/product="heme d1 biosynthesis protein NirF"
/protein_id="AAG03905.1"
/db_xref="GI:9946381"
/translation="MNLRLPLAPLLTLLAGSQOOPPLRGSGDLGVLIERADSGVQILD
GTAKTSLARVEGLDLSHSLVSRDQRYAVFGRDGLTKLDDLAQRIDKRLIOGNG
SIGGAIQDGRGLVAVSNYBEGGVKVFDSRTLELVAEIPATRLPGQDRNSRVVGLVDAP
GQRPVSLFDGSEIWIADFSQDTPHLTRFRDICKQPYDALISPDGRTYMACLFGEDG
MAQDLWHPEGRVRRVGLGYGRGOKLVPYKMPHLEGMTASDAQAPVAVGHQVVLV
DARDWKOTDAIDVAGQPVFVWFRPDQRIWYNFAYPONDQKVQVDSETHVETLRPG
PEVLHMEFSRGQDQWISVYADADQLQVMDPTRLARIGSLPARSPGSIFFSHRAQHIGL"
complement(8839..9198)
/gene="nirC"
/feature="nirC"
/feature_note="PA0517"
complement(8839..9198)
/gene="nirC"
/codon_start=1
/transl_table=11
/product="probable c-type cytochrome precursor"
/protein_id="AAG03906.1"
/db_xref="GI:9946382"
/translation="MNAPPDFRRAASHALWALALTACPLPLGLADEHDPARROAQLR
HLLLQDCSCHLRLTGGGLPALTPALRGKPRESLVATVLMGRPQTPMPWAGLLSE
DDAGWLVDRLEGEIAP"
complement(9195..9509)
/gene="nirM"
/feature="nirM"
/feature_note="PA0518"
complement(9195..9509)
/gene="nirM"
/codon_start=1
/transl_table=11
/product="cytochrome c-551 precursor"
/protein_id="AAG03907.1"
/db_xref="GI:9946383"
Query Match 27.7%; Score 39; DB 1; Length 11403;
Best Local Similarity 65.5%; Pred. No. 1.6e+02;
Matches 57; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 gccgcgagggccctggctcttgcactaccgccggtggcgccgtgacgtgtgctac 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1602 CCCACGAGGCCCTGGCGGAGTTCAGCAAGCCGCTGGCGCGCGGTGAGAACCTGAA 1661
QY 61 ccgctgagcgccgtgcggagagagcg 87
||||| ||||| ||||| ||||| |||||
Db 1662 GAGCTGACCGCTGGCTGCTGACCGC 1688
RESULT 11
AE004724
LOCUS AE004724 16903 bp DNA BCT 30-AUG-2000
DEFINITION Pseudomonas aeruginosa PA01, section 285 of 529 of the complete genome.
ACCESSION AE004724 AE004091
VERSION AE004724.1 GI:9949083
KEYWORDS
SOURCE Pseudomonas aeruginosa.
ORGANISM Pseudomonas aeruginosa.
Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
1 (bases 1 to 16903)
REFERENCE
AUTHORS Soveri,C.K., Pham,X.Q., Erwin,A.L., Mizoquchi,S.D., Warren,P.,
Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M.,
```


Mon Jun 4 11:31:04 2001

us-09-331-723a-4.rge

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2001, 13:45:48 ; Search time 19.54 Seconds
(without alignments)
281.923 Million cell updates/sec

Title: US-09-331-723A-1
Perfect score: 242
Sequence: 1 AAFAGCGENYDPPVGAVTISY ERKASGSSVPGGQIHPRTQ 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
SPREMBL15.*
1: sp.arches.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.unclassified.*
13: sp.vertebrate.*
14: sp.virus.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	%		Length	DB	ID	Description
			Match					
1	242	100.0	563	10	Q9ZTA7		Q9ZTA7	chlamydomon
2	239	98.8	563	10	Q9T0P9		Q9T0P9	chlamydomon
3	164.5	68.0	548	10	Q9SBA9		Q9SBA9	nicotiana t
4	164.5	68.0	548	10	Q9SAX9		Q9SAX9	nicotiana t
5	157.5	65.1	557	10	Q64384		Q64384	solanum tub
6	142.5	58.9	555	10	Q9SPL6		Q9SPL6	cichorium i
7	141.5	58.5	535	10	Q9M629		Q9M629	zea mays (m
8	134.5	55.6	562	10	Q9LIR8		Q9LIR8	spinacia ol
9	133.5	55.2	545	10	Q9ZS15		Q9ZS15	arabidopsis
10	57.5	23.8	284	2	Q33239		Q33239	mycobacteri
11	57.5	23.8	602	5	Q9V160		Q9V160	drosoophila
12	57.5	23.8	1317	11	Q95211		Q95211	mus musculu
13	56	23.1	350	13	Q9PWW1		Q9PWW1	lampetra re
14	56	23.1	425	3	Q12741		Q12741	blastocycladi
15	56	23.1	621	10	Q9M010		Q9M010	arabidopsis
16	56	23.1	1093	2	Q9K6B3		Q9K6B3	bacillus ha
17	55.5	22.9	1539	2	P96271		P96271	mycobacteri
18	55	22.7	343	2	Q9KZ02		Q9KZ02	streptomyce
19	55	22.7	439	2	P76130		P76130	escherichia

ALIGNMENTS

RESULT	1
ID	Q9ZTA7
CD	PRELIMINARY: PRT; 563 AA.
AC	Q9ZTA7;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE	PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
DE	PPX1.
OS	Chlamydomonas reinhardtii.
OC	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OX	Chlamydomonadaceae; Chlamydomonas.
OX	NCBI_TaxID=3055;
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=CC-621;
RX	MEDLINE=99077317; PubMed=9862501;
RX	RANDolph-Anderson B.L., Sato R., Johnson A.M., Harris E.H.,
RA	Bauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
RA	Boynton J.E.;
RT	"Isolation and characterization of a mutant protoporphyrinogen oxidase
RT	gene from Chlamydomonas reinhardtii conferring resistance to porphyric
RT	herbicides.";
RL	Plant Mol. Biol. 38:839-858(1998).
DR	EMBL: AF068635; AAC79685.1 -
DR	MENDEL: 39099; Chlre:3076; 39099.
SQ	SEQUENCE 563 AA; 5980 MW; FC5E8FFC934CAFC2 CRC64;
Query Match	100.0%; Score 242; DB 10; Length 563;
Best Local Similarity	100.0%; Pred. No. 2.5e-23;
Matches	47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AAALGSDFPYPCAVTLSPPLSAVRERKASDGSVPGFGLHPRQT 47
Dd	377 AAALGSDFPYPCAVTLSPPLSAVRERKASDGSVPGFGLHPRQT 423
RESULT	2
ID	Q9T0P9
ID	Q9T0P9
ID	Q9T0P9; PRELIMINARY: PRT; 563 AA.

DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
 DE HERBICIDE-RESISTANT PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
 GN PPX1.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-2674;
 RA Randolph-Anderson B.L., Sato R., Johnson A.M., Harris E.H.,
 RA Hauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
 RA Boynton J.E.;
 RA "Isolation and Characterization of a Mutant Protoporphyrinogen Oxidase
 RT Gene from Chlamydomonas reinhardtii Conferring Resistance to Porphyrin
 RL Herbicides".
 RL Plant Mol. Biol. 38:839-858(1998).
 DR EMBL; AF030179; AAC79630.1; -.
 KW Transit peptide.
 FT TRANSIT 1 82
 FT CHAIN 83 563
 FT POTENTIAL.
 FT HERBICIDE-RESISTANT PROTOPORPHYRINOGEN
 FT OXIDASE.
 FT VARIANT 389 389
 FT M -> V.
 FT SEQUENCE 563 AA; 59834 MW; 074584EC935CA3AF CRC64;

Query Match 98.8%; Score 239; DB 10; Length 563;
 Best Local Similarity 97.9%; Pred. No. 6.2e-23;
 Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAALGSFDYPPVGVATLSYPLSAVREERKASDGSVPGFGLHPRTQ 47
 DB 377 AAALGSFDYPPVGVATLSYPLSAVREERKASDGSVPGFGLHPRTQ 423

RESULT 3
 Q9SB14 PRELIMINARY; PRT; 548 AA.
 ID Q9SB14;
 AC Q9SB14;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE PX-1 (EC 1.3.3.4).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SR1.
 RA Horikoshi M., Mametsuka K., Hirooka T.;
 RT "The Molecular Basis of Photobleaching Herbicide Resistance in
 RL Tobacco".
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF044128; AAD02290.1; -.
 KW Oxidoreductase.
 SQ SEQUENCE 548 AA; 59241 MW; 00E1A106F299E4AF CRC64;

Query Match 68.0%; Score 164.5; DB 10; Length 548;
 Best Local Similarity 68.1%; Pred. No. 3e-13;
 Matches 32; Conservative 6; Mismatches 1; Indels 1; Gaps 1;

QY 1 AAALGSFDYPPVGVATLSYPLSAVREERKASDGSVPGFGLHPRTQ 47
 DB 364 AAALSNFYPPVGVATISYQEAIRDERLV-DGELKGFGLHPRTQ 409

RESULT 4
 Q9SA93- PRELIMINARY; PRT; 555 AA.

ID Q9SA93 PRELIMINARY; PRT; 548 AA.
 AC Q9SA93;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
 DE PLASTIDAL PROTOPORPHYRINOGEN OXIDASE.
 GN NTPOX1.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SAMSUN NN;
 RA Watanabe N., Che F., Iwano M., Takayama S., Nakano T., Yoshida S.,
 RA Isogai A.;
 RA "Molecular characterization of photomixotrophic cultured tobacco cells
 RT resistant to protoporphyrinogen oxidase-inhibiting herbicides".
 RL Plant Physiol. 118:751-758(1998).
 DR EMBL; AB020501; BAA34713.1; -.
 SQ SEQUENCE 548 AA; 59215 MW; AB51A513FD3CE7BC CRC64;

Query Match 68.0%; Score 164.5; DB 10; Length 548;
 Best Local Similarity 68.1%; Pred. No. 3e-13;
 Matches 32; Conservative 6; Mismatches 1; Indels 1; Gaps 1;

QY 1 AAALGSFDYPPVGVATLSYPLSAVREERKASDGSVPGFGLHPRTQ 47
 DB 364 AAALSNFYPPVGVATISYQEAIRDERLV-DGELKGFGLHPRTQ 409

RESULT 5
 Q64384 PRELIMINARY; PRT; 557 AA.
 ID Q64384;
 AC Q64384;
 DT 01-AUG-1998 (TEMBLrel. 07, Created)
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BINTJE;
 RA Johnston D.J., Droz E., Rochaix J.D., Malnoe P.C.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O.
 DR EMBL; AJ225107; CAAL2400.1; -.
 DR MENDEL; 29307; Soltu; 3076; 29307.
 KW Oxidoreductase.
 SQ SEQUENCE 557 AA; 60463 MW; AE2B15B17E0B89F8 CRC64;

Query Match 65.1%; Score 157.5; DB 10; Length 557;
 Best Local Similarity 66.0%; Pred. No. 2.5e-12;
 Matches 31; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 1 AAALGSFDYPPVGVATLSYPLSAVREERKASDGSVPGFGLHPRTQ 47
 DB 373 AAALSSFYPPVAAVTISYQEAIRDERLV-DGELKGFGLHPRSQ 418

RESULT 6
 Q9SPL6 PRELIMINARY; PRT; 555 AA.
 ID Q9SPL6
 AC Q9SPL6;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)


```

RESULT 11
Q9VY80 PRELIMINARY; PRT; 602 AA.
ID Q9VY80;
OC Q9VY80;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GCL1096 PROTEIN.
GN GCL1096.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscophila;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

QY 6 GSFDPVPGAVTLSYPLSAVR-EER

Query Match 23.8%; Score 57.5; DB 11; Length 1317;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 17; Conservative 3; Mismatches 11; Indels 3; Gaps 2;
QY 6 GSFDPYPVAVTSLYPLSAVR-EERKASDGSVPG 38

```

DR PFM: PF00433; pkinase_C; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR SEQUENCE 425 AA; 47035 MW; 23A3A4D538768948 CRC64;

Query Match 23.1%; Score 56; DB 3; Length 425;
Best Local Similarity 33.9%; Pred. No. 31;
Matches 19; Conservative 4; Mismatches 17; Indels 16; Gaps 2;

QY 1 AAEALGSDYPPGVAVTLSYPLSAVREERK-----ASDGSVPGFGLH 43
DB 79 ATAHTNSDYSPPAATFSAFLDAVAERRRKTTLADLELRLTGTGS---FGRVH 131

RESULT 15
Q9M010 PRELIMINARY; PRT; 621 AA.
ID Q9M010;
DC Q9M010;
DT 01-OCR-2000 (trEMBLrel. 15, Created)
DT 01-OCR-2000 (trEMBLrel. 15, Last sequence update)
DT 01-OCR-2000 (trEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 68.6 KDA PROTEIN.
GN F7A7-180.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161946; CAB82282.1; -.
KW Hypothetical protein.
SQ SEQUENCE 621 AA; 68579 MW; 259E20AAAF7ACE73F CRC64;

Query Match 23.1%; Score 56; DB 10; Length 621;
Best Local Similarity 38.9%; Pred. No. 47;
Matches 14; Conservative 7; Mismatches 13; Indels 2; Gaps 1;

QY 12 PVGAVTLSYPLSAVREERKASDGSVPGFGLHPRQTQ 47
DB 154 PVGSLEVS--LSSGGESDSSAAASHPGFSENHPDVQ 187

Search completed: June 1, 2001, 13:47:36
Job time: 108 sec

```

Mon Jun 4 11:31:03 2001

us-09-331-723a-1.rspt

Page 6

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2001, 13:45:58 ; Search time 8.09 Seconds
(without alignments)
199.012 Million cell updates/sec

Title: US-09-331-723A-1

Perfect score: 242

Sequence: 1 AALGSGFDYPPVGAFTLSY.....ERKASDGSVFCGLHPRTQ 47

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161.5	66.7	548	1 PPOC_TOBAC	O24163 nicotiana t
2	133.5	55.2	537	1 PPOC_ARATH	P5826 arabidopsis
3	60.5	25.0	403	1 GCPE_SYNT3	P73672 synechocyst
4	60.5	25.0	690	1 LIP_STAUF	O0335 staphylococ
5	58	24.0	460	1 COAA_BPFI1	O80297 bacterioph
6	57	23.6	983	1 EPA3_CHICK	P29318 gallus gall
7	57	23.6	983	1 EPA3_HUMAN	P29320 homo sapien
8	57	23.6	984	1 EPA3_RAT	O08680 rattus norv
9	56	23.1	363	1 R23A_HUMAN	P54725 homo sapien
10	55.5	22.9	450	1 YPEB_BACSU	P38490 bacillus su
11	55.5	22.9	1434	1 NOS1_HUMAN	P29475 homo sapien
12	55	22.7	163	1 YAPF_SCHPO	O09791 schizosacch
13	55	22.7	363	1 R23A_MOUSE	P54726 mus musculu
14	54	22.3	3770	1 ACVS_EMENT	P27742 emericella
15	53	21.9	288	1 PORB_METH	O27771 methanobact
16	53	21.9	983	1 EPA3_MOUSE	P29319 mus musculu
17	53	21.9	987	1 EPB2_COTJA	Q90344 coturnix co
18	53	21.9	993	1 EPB2_MOUSE	P54763 mus musculu
19	53	21.9	1004	1 EPB2_CHICK	P28693 gallus gall
20	53	21.9	1055	1 EPB2_HUMAN	P29323 homo sapien
21	53	21.9	1569	1 GLI3_XENLA	Q10660 xenopus lae
22	52.5	21.7	942	1 YD73_SCHPO	Q10328 schizosacch
23	52	21.5	234	1 CFI_VITVI	P51117 vitis vinif
24	52	21.5	323	1 JUND_CHICK	P27921 gallus gall
25	52	21.5	563	1 MDL1_PRUSE	P52706 prunus sero
26	52	21.5	787	1 SYFB_NEIMA	Q91va0 neisseria m
27	52	21.5	877	1 EPA5_MOUSE	Q60629 mus musculu
28	52	21.5	1005	1 EPA5_RAT	P54757 rattus norv
29	52	21.5	1013	1 EPA5_CHICK	P54755 gallus gall
30	52	21.5	1037	1 EPA5_HUMAN	P54756 homo sapien
31	52	21.5	3011	1 POLG_HCVH	P27958 h. genome po
32	51.5	21.3	569	1 Y4FN_RHISN	P55452 rhizobium s
33	51	21.1	387	1 SOX1_HUMAN	O00570 homo sapien

34	50.5	20.9	450	1 LAT_NOCLA	Q05174 nocardia la
35	50.5	20.9	1586	1 SN22_HUMAN	P51531 homo sapien
36	50	20.7	344	1 F16P_SYNF7	Q59943 synechococ
37	50	20.7	391	1 SOX1_MOUSE	P53783 mus musculu
38	50	20.7	605	1 WIS1_SCHPO	P33886 schizosacch
39	50	20.7	649	1 V7OK_EPMV	P20129 eggplant mo
40	50	20.7	981	1 EPA3_BRARE	O13146 brachydanio
41	50	20.7	1002	1 EPB5_CHICK	Q07497 gallus gall
42	50	20.7	1106	1 ITA7_RAT	Q63258 rattus norv
43	49.5	20.5	229	1 CMF3_BACSU	P39147 bacillus su
44	49.5	20.5	341	1 JUND_MOUSE	P15066 mus musculu
45	49.5	20.5	341	1 JUND_RAT	P52909 rattus norv

ALIGNMENTS

RESULT	1
PPOC_TOBAC	
ID	PPOC_TOBAC
AC	O24163; STANDARD; PRT: 548 AA.
DT	15-DEC-1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	PROTOPORPHYRIN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO I)
DE	(PROTOPORPHYRIN OXIDASE ISOZYME I) (PPX I).
GN	PPXI.
OS	Nicotiana tabacum (Common tobacco).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC	Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC	Solanales; Solanaceae; Nicotiana.
OX	NCBI_TaxID=4097;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. SR1;
RX	MEDLINE=97385200; PubMed=9238074;
RA	Lermontova I., Kruse E., Mock H.-P., Grimm B.;
RT	"Cloning and characterization of a plastid and a mitochondrial
RT	isoform of tobacco protoporphyrinogen IX oxidase.";
RL	Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).
CC	!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRIN
CC	IX TO FORM PROTOPORPHYRIN IX.
CC	!- CATALYTIC ACTIVITY: PROTOPORPHYRIN-IX + O(2) = PROTOPORPHYRIN-
CC	IX + H(2)O(2).
CC	!- COFACTOR: (2).
CC	!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS, AND
CC	CHLOROPHYLL SYNTHESIS.
CC	!- SUBUNIT: HOMODIMER; CONTAINS ONE FAD PER HOMODIMER (BY
CC	SIMILARITY).
CC	!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC	!- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES.
CC	DECREASED EXPRESSION IN OLDEST LEAVES. EXPRESSED AT VERY LOW LEVEL
CC	IN ROOTS.
CC	!- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL
CC	EXPRESSION IN THE DARK PERIOD.
CC	!- SIMILARITY: BELONGS TO THE PROTOPORPHYRIN OXIDASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; Y13465; CAA73865.1; -
DR	Porphylin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW	Transit peptide; Flavoprotein; FAD; Chloroplast;
KW	Chlorophyll biosynthesis.
FT	CHAIN 1 50 CHLOROPLAST (POTENTIAL).
FT	TRANSIT 51 58 PROTOPORPHYRIN OXIDASE.
FT	NP_BIND 78 83 FAD (POTENTIAL).
FT	SEQUENCE 548 AA; 59230 MW; 66892E78F8A3E30 CRC64;
SQ	

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
DR EMBL; D90908; BAA17717.1; -
SQ SEQUENCE 403 AA; 44176 MW; 23948D183237DDB8 CRC64;

Query Match 25.0%; Score 60.5; DB 1; Length 403;
Best Local Similarity 43.2%; Pred. No. 1.2;
Matches 16; Conservative 6; Mismatches 10; Indels 5; Gaps 2;

QY 12 PVGAVTL--SYPL---SAVREERKASDGVPGFGQLH 43

Db 27 PVGAVTGGHVVQVQSMINEDLVDGVSAGIRRLH 63

RESULT 4

ID LIP_STAAT STANDARD; PRT; 690 AA.

AC P10335;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE LIPASE PRECURSOR (EC 3.1.1.3) (GLYCEROL ESTER HYDROLASE).

GN GER.

OS Staphylococcus aureus.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; staphylococcus.

OX NCBI_TaxID=1280;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86195821; PubMed=3009394;

RA Lee C.Y., Iandolo J.J.;

RT "Lysogenic conversion of staphylococcal lipase is caused by insertion
RT of the bacteriophage L54a genome into the lipase structural gene.;"

RL J. Bacteriol. 156:385-391(1986).

RN [2]

RP PROCESSING, AND SEQUENCE OF 296-307.

RC STRAIN=105;

RX MEDLINE=92193269; PubMed=1548232;

RA Roloff J., Normark S.;

RT "In vivo processing of Staphylococcus aureus lipase.;"

RL J. Bacteriol. 174:1844-1847(1992).

CC -!- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
CC A FATTY ACID ANION.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- MISCELLANEOUS: THE EXPRESSION OF STAPHYLOCOCCUS LIPASE IS

CC NEGATIVELY REGULATED BY BACTERIOPHAGE LYSOGENIZATION (LIPASE

CC CONVERSION). PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,

CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
DR EMBL; M12715; AAA26633.1; -
DR PIR; A24545; A24545.

DR LipoPro; IPR000734; -

DR PROSITE; PS00120; LIPASE_SER; 1.

KW Hydrolase; Lipid degradation; Zymogen; Signal.

FT SIGNAL 1 37

FT PROPEP 38 295

FT CHAIN 296 690

FT LIPASE.

FT DOMAIN 311 690 HYDROPHOBIC.
FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 645 645 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 690 AA; 76388 MW; 8E8E3654D0E01A3B CRC64;

Query Match 25.0%; Score 60.5; DB 1; Length 690;

Best Local Similarity 33.3%; Pred. No. 2.2;

Matches 15; Conservative 6; Mismatches 19; Indels 5; Gaps 1;

QY 8 FDYPPVGAUTLSYPLSAV-----REERKASDGVPGFGQLHPRQ 47

Db 573 YENPLDGLFMATTSRIIHDAREWRKNDGVVPISSLHPSNQ 617

RESULT 5

CORA_BP1F1

ID COAA_BP1F1 STANDARD; PRT; 460 AA.

AC O80297;

DT 13-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE COAT PROTEIN A PRECURSOR (ATTACHMENT PROTEIN).

GN III OR 3.

OS Bacteriophage If1.

OC Viruses; ssDNA viruses; Inoviridae; Inovirus.

OX NCBI_TaxID=10868;

RN [1]

RP SEQUENCE FROM N.A.

RA Hill D.F., Hughes G., McNaughton J.C., Stockwell P.A., Petersen G.B.;

RT "DNA sequence of the filamentous coliphage If1.;"

RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: COAT PROTEIN A IS NECESSARY FOR ADSORPTION OF THE VIRION

CC ONTO THE PILIUS OF THE HOST CELL (BY SIMILARITY).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
DR EMBL; U02303; AAC62155.1; -

DR Phage recognition; Coat protein; Signal.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 460 COAT PROTEIN A.

SQ SEQUENCE 460 AA; 48790 MW; 9AD651968C715AB2 CRC64;

Query Match 24.0%; Score 58; DB 1; Length 460;

Best Local Similarity 43.6%; Pred. No. 3.1;

Matches 17; Conservative 6; Mismatches 10; Indels 6; Gaps 2;

QY 6 GSFDPPVG-----AVTLYSPLSAVREERKAS-DGSVPG 38

Db 305 GDFYVKMANAKDALTESFDLSAQADTGASLDGVSQG 343

RESULT 6

EP3_CHICK

ID EPA3_CHICK STANDARD; PRT; 983 AA.

AC P29318;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN

DE KINASE RECEPTOR ETK1) (CEK4).

GN EPA3 OR ETK1 OR CEK4.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.


```

DR InterPro: IPR001090; -
DR InterPro: IPR001245; -
DR InterPro: IPR001426; -
DR InterPro: IPR001660; -
DR InterPro: IPR001777; -
DR Pfam: PF01404; EPH1bd; 1.
DR Pfam: PF00536; SAM; 1.
DR Pfam: PF00041; fn3; 2.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00014; FNTYPEII.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR PROSITE: PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
KW Transferase: Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 20
FT CHAIN 21 983
FT DOMAIN 21 541
FT TRANSMEM 542 565
FT DOMAIN 566 983
FT DOMAIN 189 322
FT DOMAIN 323 432
FT DOMAIN 433 530
FT DOMAIN 621 882
FT DOMAIN 909 983
FT SITE 981 983
FT NP_BIND 627 635
FT BINDING 633 653
FT ACT_SITE 746 746
FT MOD_RES 596 596
FT MOD_RES 602 602
FT MOD_RES 779 779
FT CARBOHYD 232 232
FT CARBOHYD 337 337
FT CARBOHYD 391 391
FT CARBOHYD 404 404
FT CARBOHYD 453 453
FT CONFLICT 507 507
FT CONFLICT 724 724
SQ SEQUENCE 983 AA; 110086 MW; B8D900FA80FF5121 CRC64;

Query Match 23.6%; Score 57; DB 1; Length 983;
Best Local Similarity 34.8%; Pred. No. 9.9;
Matches 16; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY 2 AELGSGFDYPPV---GAVILSYPLSAVREERKASDGSVFGQLH 43
DB 671 ASIMGQDFHNIIRLEGVTKRPVMIVTEYME--NGSLDSFLRKH 714

RESULT 8
ID EPA3_RAT STANDARD; PRT; 984 AA.
AC 008680;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
GN KINASE RECEPTOR REK4) (TYRO-4).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA MEDLINE=98120505; PubMed=9458884;

```

```

RA Li Y.Y., McTiernan C.F., Feldman A.M.;
RT "IL-1 beta alters the expression of the receptor tyrosine kinase gene
RL i-EphA3 in neonatal rat cardiomyocytes.";
CC Am. J. Physiol. 274:H331-H341(1998).
CC -1- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC EPHRIN-A2, -A3, -A4 AND -A5.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN THE HEART, BRAIN, AND LUNG.
CC -1- INDUCTION: DOWNREGULATED BY INTERLEUKINE-1BETA IN NEONATAL CARDIAC
CC MYOCYTES.
CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN
CC RECEPTOR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U69278; AAC06273.1; -
CC HSSP: P16109; 1FSB.
CC InterPro: IPR000561; -
CC InterPro: IPR000719; -
CC InterPro: IPR001090; -
CC InterPro: IPR001245; -
CC InterPro: IPR001426; -
CC InterPro: IPR001660; -
CC InterPro: IPR001777; -
CC Pfam: PF01404; EPH1bd; 1.
CC Pfam: PF00536; SAM; 1.
CC Pfam: PF00041; fn3; 2.
CC Pfam: PF00069; pkinase; 1.
CC PRINTS: PR00014; FNTYPEII.
CC PRINTS: PR00109; TYRKINASE.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00790; RECEPTOR_TYR_KIN_V_1; 1.
CC PROSITE: PS00791; RECEPTOR_TYR_KIN_V_2; 1.
CC PROSITE: PS01186; EGF_2; UNKNOWN_1.
KW Transferase: Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 20
FT CHAIN 21 984
FT DOMAIN 21 541
FT TRANSMEM 542 565
FT DOMAIN 566 984
FT DOMAIN 189 322
FT DOMAIN 323 431
FT DOMAIN 436 528
FT DOMAIN 622 883
FT DOMAIN 910 984
FT SITE 982 984
FT NP_BIND 628 636
FT BINDING 634 654
FT ACT_SITE 747 747
FT MOD_RES 597 597
FT MOD_RES 603 603
FT MOD_RES 780 780
FT CARBOHYD 232 232
FT CARBOHYD 337 337
FT CARBOHYD 391 391
FT CARBOHYD 404 404
FT CARBOHYD 493 493
SQ SEQUENCE 984 AA; 110227 MW; F170C49312F7A0AB CRC64;

```

Query Match 23.6%; Score 57; DB 1; Length 984;
Best Local Similarity 34.8%; Pred. No. 9.9;
Matches 16; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY 2 AALGSGFDYPPV-----GAVTLSPYLSAVREERKASDGSVPFGQLH 43
DB 572 ASIMQGDHPNIRLEGGVTKSPVMIVTEYME--NGSLDSFLRXH 715

RESULT 9
R23A_HUMAN STANDARD; PRT; 363 AA.
AC P54725;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UV EXCISION REPAIR PROTEIN RAD23 HOMOLOG A (HHR23A).
GN RAD23A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
GN [1]
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94222030; PubMed=8168482;
RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ui M.,
RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,
RA Hoeljmakers J.H.J., Hanaoka F.;
RT "Purification and cloning of a nucleotide excision repair complex
involving the xeroderma pigmentosum group C protein and a human
homologue of yeast RAD23";
RT EMBO J. 13:1831-1843(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Lamerdin J., McCreedy P., Stillwagen S., Ramirez M., Carrano A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP STRUCTURE BY NMR OF 319-363.
RX MEDLINE=99061330; PubMed=9846873;
RA Dieckmann T., Withers-Ward E.S., Jarosinski M.A., Liu C.F.,
RA Chen I.S.Y., Feigon J.;
RT "Structure of a human DNA repair protein UBA domain that interacts
with HIV-1 Vpr";
RL Nat. Struct. Biol. 5:1042-1047(1998).
RN [4]
RP STRUCTURE BY NMR OF 319-363.
RX MEDLINE=20541363; PubMed=11087358;
RA Withers-Ward E.S., Mueller T.D., Chen I.S.Y., Feigon J.;
RT "Biochemical and structural analysis of the interaction between the
UBA(2) domain of the DNA repair protein HHR23A and HIV-1 Vpr";
RL Biochemistry 39:14103-14112(2000).
CC 1- FUNCTION: COULD BE INVOLVED IN DNA EXCISION REPAIR.
CC 1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC 1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.
CC 1- SIMILARITY: CONTAINS 2 UBA DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D21235; BAA04767.1; -;
DR EMBL: AD000092; AAB51177.1; -;
DR FDB: IDV0; 11-FEB-00.
DR MIM: 600061; -;
DR InterPro: IPR000449; -;
DR InterPro: IPR000626; -;
DR Pfam: PF00627; UBA; 2.
DR Pfam: PF00240; ubiquitin; 1.
DR PROSITE: PS00053; UBIQUITIN_2; 1.

KW DNA damage; DNA repair; Nuclear protein; 3D-structure.
FT DOMAIN 1 81 UBIQUITIN-LIKE.
SQ SEQUENCE 363 AA; 39609 MW; C4E47E9313BB47B5 CRC64;

Query Match 23.1%; Score 56; DB 1; Length 363;
Best Local Similarity 35.5%; Pred. No. 4.4;
Matches 11; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 7 SFDYPPVGAVTLSYPLSAVREERKASDGSVP 37
DB 100 STSFPAPTSGMSHPAPAREDKSPSEESAP 130

RESULT 10
YPEB_BACSU STANDARD; PRT; 450 AA.
AC P38490; P40774;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 51.2 KDA PROTEIN IN RECO-CMK INTERGENIC REGION (PSPA12).
GN YPEB OR JOEB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168 / MARBURG;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
the serA and kog loci cloned in a yeast artificial chromosome";
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE OF 220-450 FROM N.A.
RX STRAIN=168 / MARBURG;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-331 FROM N.A.
RX STRAIN=168;
RX MEDLINE=96427343; PubMed=8830707;
RA Moriyama K., Hattori A., Miyata S., Kudoh S., Makino S.;
RT "A gene (slsB) encoding a spore cortex-lytic enzyme from Bacillus
subtilis and response of the enzyme to L-alanine-mediated
germination";
RL J. Bacteriol. 178:6059-6063(1996).
RN [4]
RP SEQUENCE OF 1-27 FROM N.A.
RX MEDLINE=89108019; PubMed=3145906;
RA Smith H., de Jong A., Bron S., Venema G.;
RT "Characterization of signal-sequence-coding regions selected from the
Bacillus subtilis chromosome";
RL Gene 70:351-361(1988).
CC 1- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 17.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L47648; AAC83958.1; -;
DR EMBL: U11687; AAA85146.1; -;
DR EMBL: D79978; BAA11474.1; -;
DR EMBL: M22904; AAA22820.1; ALT_FRAME.

DR EMBL; Z99115; CAB14208.1; -;
DR EMBL; Z99116; CAB14224.1; -;
DR Subtilist; BG11001; ypeB.
KW Hypothetical protein.
SQ SEQUENCE 450 AA; 51185 MW; 8B4A7B479C088E6B CRC64;

Query Match 22.9%; Score 55.5; DB 1; Length 450;
Best Local Similarity 37.8%; Pred. No. 6.5;
Matches 14; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

OY 3 EALGSDYPPVGAFTLSYPSLAVREERKASDGVPGF 39
: | | | | | | | | | | | | | | | | | | | |
DB 324 DKIGVSYVPVENKVRMYP-EAIRMKVALDDGEVGF 359

RESULT 11
NOS1_HUMAN
ID NOS1_HUMAN STANDARD; PRT; 1434 AA.
AC P29475;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NITRIC-OXIDE SYNTHASE, BRAIN (EC 1.14.13.39) (NOS, TYPE I) (NEURONAL
DE NOS) (N-NOS) (NNOS) (CONSTITUTIVE NOS) (NC-NOS) (BNOS).
GN NOS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM N-NOS-1).
RX MEDLINE-95105197; PubMed=7528745;
RA Hall A.V., Antoniou H., Wang Y., Cheung A.H., Arbus A.M., Olson S.L.,
RA Lu W.C., Kau C.-L., Marsden P.A.;
RT "Structural organization of the human neuronal nitric oxide synthase
RT gene (NOS1).";
RL J. Biol. Chem. 269:33082-33090(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS N-NOS-1 AND N-NOS-2).
RC TISSUE=Cerebellum;
RX MEDLINE-94267447; PubMed=7515942;
RA Fujisawa H., Ogura T., Kurashima Y., Yokoyama T., Yamashita J.,
RA Esumi H.;
RT "Expression of two types of nitric oxide synthase mRNA in human
RT neuroblastoma cell lines.";
RL J. Neurochem. 63:140-145(1994).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM N-NOS-1).
RC TISSUE=Brain;
RX MEDLINE-93131039; PubMed=7678401;
RA Nakane M., Schmidt H.H.W., Pollock J.S., Foerstermann U., Murad F.;
RT "Cloned human brain nitric oxide synthase is highly expressed in
RT skeletal muscle.";
RL FEBS Lett. 316:175-180(1993).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM N-NOS-1).
RC TISSUE=Retina;
RX MEDLINE-97034095; PubMed=8879752;
RA Park C.-S., Gnanou C., Park R., Krishna G.;
RT "Neuronal isoform of nitric oxide synthase is expressed at low levels
RT in human retina.";
RL Cell. Mol. Neurobiol. 16:499-515(1996).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS TN-NOS; TN-NOSB AND TEX2 INSERTION).
RC TISSUE=Testis;
RX MEDLINE-97269048; PubMed=9111048;
RA Wang Y., Golligorsky M.S., Lin M., Wilcox J.N., Marsden P.A.;
RT "A novel, testis-specific mRNA transcript encoding an NH2-terminal
RT truncated nitric-oxide synthase.";
RL J. Biol. Chem. 272:11392-11401(1997).
CC -1- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE
WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY. IN THE BRAIN AND

PERIPHERAL NERVOUS SYSTEM, NO DISPLAYS MANY PROPERTIES OF A
NEUROTRANSMITTER.
-1- CATALYTIC ACTIVITY: L-ARGININE + N NADPH + M O(2) = CITRULLINE +
NITRIC OXIDE + N NADP(+).
-1- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES
TETRAHYDROBIOTIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF
THE ENZYME.
-1- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN. INHIBITED BY
N-NOS-INHIBITING PROTEIN (PIN) WHICH MAY PREVENT THE DIMERIZATION
OF THE PROTEIN.
-1- SUBUNIT: HOMODIMER.
-1- SUBCELLULAR LOCATION: IN SKELETAL MUSCLE, IT IS LOCALIZED BENEATH
THE SARCOLEMA OF FAST-TWITCH MUSCLE FIBER BY ASSOCIATING WITH THE
DYSTROPHIN GLYCOPROTEIN COMPLEX.
-1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; N-NOS-1 (SHOWN HERE), N-NOS-
2, TN-NOS/TN-NOSB AND TEX2 INSERTION; ARE PRODUCED BY ALTERNATIVE
SPLICING. TN-NOS/TN-NOSB ARE PRODUCED BY DIFFERENT ALTERNATIVE
(TN-NOS) OR TEX2 (TN-NOSB) LEADING TO A NH2-TRUNCATED PROTEIN
WHICH POSSESSES ENZYMIC ACTIVITY COMPARABLE TO THAT OF N-NOS-1.
OF THE TEX2 EXON BETWEEN EXONS 3 AND 4 OF THE N-NOS-1 ISOFORM,
LEADING TO A FRAMESHIFT AND A PREMATURE STOP CODON.
-1- TISSUE SPECIFICITY: THE N-NOS-1 ISOFORM IS UBICUITOUSLY EXPRESSED:
DETECTED IN SKELETAL MUSCLE AND BRAIN, ALSO IN TESTIS, LUNG AND
KIDNEY, AND AT LOW LEVELS IN HEART, ADRENAL GLAND AND RETINA. NOT
DETECTED IN THE PLATELETS. THE TN-NOS ISOFORM IS EXPRESSED ONLY IN
TESTIS. THE TN-NOSB ISOFORM IS DETECTED AT LOW LEVELS ONLY IN THE
TESTIS. THE TEX2 INSERTION ISOFORM IS DETECTED IN TESTIS, SKELETAL
MUSCLE, LUNG, AND KIDNEY, AT LOW LEVELS IN THE BRAIN, BUT NOT IN
THE HEART AND ADRENAL GLAND.
-1- DOMAIN: THE PDZ DOMAIN IN THE N-TERMINAL PART OF THE NEURONAL
ISOFORM PARTICIPATES IN PROTEIN-PROTEIN INTERACTION, AND IS
RESPONSIBLE FOR TARGETING NNOS TO SYNAPTIC MEMBRANES IN MUSCLES
(BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE NOS FAMILY.
-1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U17327; AAA62405.1; -;
EMBL; U17326; AAB60654.1; ALT_SEQ.
EMBL; U17299; AAB60654.1; JOINED.
EMBL; U17300; AAB60654.1; JOINED.
EMBL; U17301; AAB60654.1; JOINED.
EMBL; U17302; AAB60654.1; JOINED.
EMBL; U17303; AAB60654.1; JOINED.
EMBL; U17304; AAB60654.1; JOINED.
EMBL; U17305; AAB60654.1; JOINED.
EMBL; U17307; AAB60654.1; JOINED.
EMBL; U17308; AAB60654.1; JOINED.
EMBL; U17309; AAB60654.1; JOINED.
EMBL; U17310; AAB60654.1; JOINED.
EMBL; U17311; AAB60654.1; JOINED.
EMBL; U17312; AAB60654.1; JOINED.
EMBL; U17313; AAB60654.1; JOINED.
EMBL; U17314; AAB60654.1; JOINED.
EMBL; U17315; AAB60654.1; JOINED.
EMBL; U17316; AAB60654.1; JOINED.
EMBL; U17317; AAB60654.1; JOINED.
EMBL; U17318; AAB60654.1; JOINED.
EMBL; U17319; AAB60654.1; JOINED.
EMBL; U17320; AAB60654.1; JOINED.
EMBL; U17321; AAB60654.1; JOINED.
EMBL; U17322; AAB60654.1; JOINED.
EMBL; U17323; AAB60654.1; JOINED.
EMBL; U17324; AAB60654.1; JOINED.

DR EMBL: U17325; AAB60654.1; JOINED.
 DR EMBL: D16408; BAA03895.1; -.
 DR EMBL: U02881; AAA36376.1; -.
 DR EMBL: U31466; AAB49040.1; -.
 DR EMBL: U66362; -; NOT_ANNOTATED_CDS.
 DR PIR: S28878; S28878.
 DR HSSP: P00388; LAMO.
 DR MIM: 163731; -.
 DR InterPro: IPR001094; -.
 DR InterPro: IPR001433; -.
 DR InterPro: IPR001478; -.
 DR InterPro: IPR001709; -.
 DR InterPro: IPR003097; -.
 DR Pfam: PF00667; FAD_binding; 1.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF00175; oxidored_fad; 1.
 DR PRINTS: PR00369; FLAVODOXIN.
 DR PRINTS: PR00371; FPNCR.
 DR PROSITE: PS0106; PDZ; 1.
 DR Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme;
 KW Alternative splicing; Multigene family.
 FT DOMAIN 17 99 PDZ
 FT BINDING 163 245 NNOS-INHIBITING PROTEIN (PIN)-BINDING.
 FT BINDING 420 420 HEME (BY SIMILARITY).
 FT DOMAIN 730 750 CALMODULIN-BINDING (POTENTIAL).
 FT NP_BIND 886 917 FMN (PYRIMIDINE PART) (BY SIMILARITY).
 FT NP_BIND 1032 1043 FAD (ADP PART) (BY SIMILARITY).
 FT NP_BIND 1175 1185 FAD (FLAVIN PART) (BY SIMILARITY).
 FT NP_BIND 1250 1268 NADP (RIBOSE PART) (BY SIMILARITY).
 FT NP_BIND 1348 1363 NADP (ADP PART) (BY SIMILARITY).
 FT VARSPLIC 1 336 MISSING (IN ISOFORM TN-NOS/TN-NOSB).
 FT VARSPLIC 285 407 PPTSGKSPKNGSKSPKPRFLKYNKMETVLTDLHLKS
 FT TLENGCTEYICGSIHPSOHARPEQVTKGOLFPLAKEF
 FT IDQYSSIKBEGSKAHMERLEVAKEIDTSTYQLKDTLI
 FT -> MKRLRIETGFGVGRSHNHPQSPQKMAAPPSVH
 FT ASSRGTGRULVSUPSTLRHAKWRALSTSAWAPSCILL
 FT SMQGLTSKQKSDSSLSKLLINTIHLQDLAPKPTWKG
 FT WKR (IN ISOFORM TEX2 INSERTION).
 FT MISSING (IN ISOFORM TEX2 INSERTION).
 FT MISSING (IN ISOFORM N-NOS-2).
 FT K -> E (IN REF. 4).
 FT LAPRPPG -> WQAPR (IN REF. 3 AND 4).
 FT QP -> HR (IN REF. 3).
 FT V -> L (IN REF. 3).
 FT G -> A (IN REF. 3).
 FT Y -> I (IN REF. 3).
 FT VARSPLIC 408 1434
 FT VARSPLIC 509 613
 FT CONFLICT 131 131
 FT CONFLICT 178 184
 FT CONFLICT 492 493
 FT CONFLICT 549 549
 FT CONFLICT 563 563
 FT CONFLICT 1407 1407
 FT SEQUENCE 1434 AA; 160970 MW; 99235793B53BF37 CRC64;

Query Match 22.9%; Score 55.5; DB 1; Length 1434;
 Best Local Similarity 47.1%; Pred. No. 24;
 Matches 16; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

QY 11 PPGVAVTSLYPLSAVREERKASDG-SVPGFGQLH 43
 Db 128 PPTKAVDLSHQPPAGKEQPLAVDASGFGNGPQH 161

RESULT 12

ID Y9F_SCHPO STANDARD; PRT; 163 AA.
 AC Q09791;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HYPOTHETICAL 18.1 KDA PROTEIN C1366.15C IN CHROMOSOME 1.
 GN SPAC1366.15C OR SPAC24B11.04C
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Odell C., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z54308; CAA91108.1; -.
 DR EMBL: Z67757; CAA91769.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 163 AA; 18051 MW; F2575B332A276D32 CRC64;

Query Match 22.7%; Score 55; DB 1; Length 163;
 Best Local Similarity 42.4%; Pred. No. 24;
 Matches 14; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

QY 13 VGAVTSLYPLSAVREERKASDGSVPGFGQLH 43
 Db 33 LGKVLASYNDKAVEEDTLKSKSTGSLPSGQGVH 65

RESULT 13

ID R23A_MOUSE STANDARD; PRT; 363 AA.
 AC P54726;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE UV EXCISION REPAIR PROTEIN RAD23 HOMOLOG A (MHR23A).
 GN RAD23A OR MHR23A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Testis;
 RX MEDLINE=96403997; Pubmed=8808275;
 RA van der Spek P.J., Visser C.E., Hanaoka F., Smit B.,
 RA Hagemeijer A., Bootsma D., Hoeljmakers J.H.J.;
 RT "Cloning, comparative mapping, and RNA expression of the mouse
 RT homologues of the Saccharomyces cerevisiae nucleotide excision repair
 RT gene RAD23.";
 RL Genomics 31:20-27(1996).
 CC -!- FUNCTION: COULD BE INVOLVED IN DNA EXCISION REPAIR.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -!- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 UBA DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92410; CAA63145.1; -.
 DR HSSP: P54725; LUBA.
 DR MGD: MGI-105126; Rad23a.
 DR InterPro: IPR000449; -.
 DR InterPro: IPR000626; -.
 DR Pfam: PF00627; UBA; 2.
 DR Pfam: PF00240; ubiquitin; 1.
 DR PROSITE: PSS0053; UBIQUITIN_2; 1.
 DR DNA damage; DNA repair; Nuclear protein.

QY 11 PPVGAVTLSYPLSAVREERKASDGSVPFGGOLH 43
Db 173 PYVATASISYPEDFMEKVRKAKETDGPAYIHLH 205

Search completed: June 1, 2001, 13:47:48
Job time: 110 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2001, 13:45:18 ; Search time 12.98 seconds
(without alignments)
248.843 Million cell updates/sec

Title: US-09-331-723A-1
Perfect score: 242
Sequence: 1 AAALGSFDYPPVGAVALTSY.....ERKASDGSVPFGQLHPTQ 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-67:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161.5	66.7	548	2 T04058	protoporphyrinogen
2	157.5	65.1	557	2 T07116	protoporphyrinogen
3	133.5	55.2	545	2 T02005	protoporphyrinogen
4	60.5	25.0	403	2 S77159	gcpg protein - Syn
5	60.5	25.0	690	2 A24545	triacylglycerol li
6	57.5	23.8	284	2 E70506	hypothetical prote
7	57.5	23.8	1317	2 T03748	apoptosis associat
8	57	23.6	983	2 B45583	receptor tyrosine
9	57	23.6	983	2 A38224	protein-tyrosine k
10	56	23.1	363	2 S44443	RAD23 protein homo
11	56	23.1	425	2 S41099	protein kinase (EC
12	56	23.1	621	2 T48187	hypothetical prote
13	56	23.1	1093	2 H84126	cation efflux syst
14	55.5	22.9	450	2 E69934	conserved hypotet
15	55.5	22.9	1433	2 G01946	nitric-oxide synth
16	55.5	22.9	1539	2 G06300	probable ctpH prot
17	55	22.7	163	2 S62444	conserved hypotet
18	55	22.7	439	2 F64902	hypothetical prote
19	54.5	22.5	152	2 T51796	hypothetical prote
20	54	22.3	143	2 G72492	hypothetical prote
21	54	22.3	283	2 S75639	hypothetical prote
22	54	22.3	3770	2 A40889	delta-(L-alpha-ami
23	53	21.9	190	2 B81042	conserved hypotet
24	53	21.9	288	2 D69099	pyruvate synthase
25	53	21.9	303	2 S76535	probable phosphoe
26	53	21.9	478	2 A57174	protein-tyrosine k
27	53	21.9	480	2 I48760	protein-tyrosine k
28	53	21.9	970	2 I78842	receptor protein-t
29	53	21.9	983	2 A45583	receptor tyrosine

ALIGNMENTS

RESULT 1

T04058
protoporphyrinogen oxidase (EC 1.3.3.4) IX precursor, chloroplast - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T04058
R:Lernontova, I.; Kruse, E.; Mock, H.P.; Grimm, B.
Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997
A:Title: Cloning and characterization of a plastidal and a mitochondrial isoform of t
A:Reference number: Z15186; MUID:97385200
A:Accession: T04058
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-548 <LER>
A:Cross-references: EMBL:Y13465; NID:g2370332; PIDN:CAA73865.1; PID:g2370333
A:Experimental source: strain SRI
C:Genetics:
A:Gene: ppxi
A:Function:
A:Pathway: tetrapyrrole synthesis
A:Keywords: chloroplast; oxidoreductase
F:1-50/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:51-548/Product: protoporphyrinogen oxidase IX #status predicted <MAT>

Query Match 66.7%; Score 161.5; DB 2; Length 548;
Best Local Similarity 66.0%; Pred. No. 8.5e-14;
Matches 31; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 1 AAALGSFDYPPVGAVALTSYPLSAVREERKASDGSVPFGQLHPTQ 47
DB 364 AAALSNFYPPVGAVALTITTPQAIKIDRLV-DGELKGFGLHPTQ 409

RESULT 2

T07116
protoporphyrinogen oxidase (EC 1.3.3.4) IX, chloroplast - potato
C:Species: Solanum tuberosum (potato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
C:Accession: T07116
R:Johnston, D.J.
submitted to the EMBL Data Library, April 1998
A:Reference number: Z15932
A:Accession: T07116
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-557 <JOH>
A:Cross-references: EMBL:AJ225107; NID:e1289157; PIDN:CAA12400.1; PID:e1289158
A:Experimental source: cv. Bintje
C:Genetics:
A:Genome: nuclear
C:Function:

protein kinase (EC 2.7.1.37), cAMP-dependent, catalytic chain C - fungus (Blastocladiella)
 C:Species: Blastocladiella emersonii
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 28-May-1999
 C:Accession: S41099; S77889; S77890
 R:Ranço de Oliveira, J.C.; Cantisani Borges, A.C.; do Valle Marques, M.; Lopes Gomes, S.
 Eur. J. Biochem. 219, 555-562, 1994
 A:Title: Cloning and characterization of the gene for the catalytic subunit of cAMP-depe
 A:Reference number: S41099; MUID:94139736
 A:Accession: S41099
 A:Molecule type: DNA
 A:Residues: 1-425 <FRA>
 A:Cross-references: GB:L17008; NID:g304272; PIDN:AAA20074.1; PID:g304273
 A:Accession: S77889
 A:Molecule type: mRNA
 A:Residues: 22-425 <FRB>
 A:Cross-references: GB:M81709; GB:L17038; NID:g507140; PIDN:AAA19440.1; PID:g507141
 A:Accession: S77890
 A:Molecule type: protein
 A:Residues: 2-16 <FRC>
 C:Genetics:
 A:Introns: 209/3; 243/3; 315/1
 A:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP; cAMP binding; magnesium; phosphoprotein; phosphotransferase; serine/thr
 F:2-425/Product: protein kinase, cAMP-dependent, catalytic chain C #status experimental
 F:114-370/Domains: protein kinase homology <KIN>
 F:122-130/Region: protein kinase ATP-binding motif
 F:127,128,194,200,243,256/Binding site: Mg-ATP (Phe, Gly, Glu, Glu, Thr) #status pr
 F:145,164,239,241/Active site: Lys, Glu, Asp, Lys #status predicted
 F:244,257/Binding site: magnesium (Asn, Asp) #status predicted
 F:270/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 23.1%; Score 56; DB 2; Length 425;
 Best Local Similarity 33.9%; Pred. No. 10;
 Matches 19; Conservative 4; Mismatches 17; Indels 16; Gaps 2;
 QY 1 AAELGSGFDYPPGVATLSYPLSAVREK-----ASDGSVPGFGLH 43
 DB 79 ATAHTNSDYSPPAATPSAPLDAVARRRKTTLADLELRQTLGTGS---FGRVH 131

RESULT 12
 T48187
 hypothetical protein F7A7.180 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48187
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
 ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: 224487
 A:Accession: T48187
 A:Molecule type: DNA
 A:Residues: 1-621 <BEV>
 A:Cross-references: EMBL:AL161946
 A:Experimental source: cultivar Columbia; BAC clone F7A7
 C:Genetics:
 A:Map position: 5
 A:Introns: 65/3; 263/3; 289/3; 403/1; 469/3; 495/2; 529/2; 592/3
 A:Note: F7A7.180

Query Match 23.1%; Score 56; DB 2; Length 621;
 Best Local Similarity 38.9%; Pred. No. 16;
 Matches 14; Conservative 7; Mismatches 13; Indels 2; Gaps 1;
 QY 12 PVGAVTSLYPLSAVREKASDGSVPGFGLHPRTQ 47
 DB 154 PVGSLEVS--LSSGSDSSAAASHPGFSENHPDVQ 187

RESULT 13

H84126
 cation efflux system BH3816 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
 C:Accession: H84126
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
 A:Reference number: A83650; MUID:20263314
 A:Accession: H84126
 A:Molecule type: DNA
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1093 <STO>
 A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07535.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3816

Query Match 23.1%; Score 56; DB 2; Length 1093;
 Best Local Similarity 58.8%; Pred. No. 30;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 11 PPVGAVTSLYPLSAVRE 27
 DB 40 PPVGAVTASYPGASAE 56

RESULT 14
 E69934
 conserved hypothetical protein ypeB - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 01-Dec-2000
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A: Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieser, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanl
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
 keuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
 A:Reference number: A69580; MUID:98044033
 A:Accession: E69934
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-450 <KUN>
 A:Cross-references: GB:Z99115; GB:Z99116; GB:AL009126; NID:g2634723; PIDN:CAB14224.1;
 A:Experimental source: strain 168
 R:Moriyama, R.; Hattori, A.; Miyata, S.; Kudoh, S.; Makino, S.
 J. Bacteriol. 178, 6059-6063, 1996
 A:Title: A gene (steB) encoding a spore cortex-lytic enzyme from Bacillus subtilis an
 A:Reference number: Z22836; MUID:96427343
 A:Accession: T44771
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-331 <MOR>
 A:Cross-references: EMBL:D79978; NID:gl688021; PIDN:BAAL1474.1; PID:gl688024
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ypeB

Query Match 22.9%; Score 55.5; DB 2; Length 450;
 Best Local Similarity 37.8%; Pred. No. 13;
 Matches 14; Conservative 4; Mismatches 18; Indels 1; Gaps 1;
 QY 12 PVGAVTSLYPLSAVREKASDGSVPGFGLHPRTQ 47
 DB 154 PVGSLEVS--LSSGSDSSAAASHPGFSENHPDVQ 187

RESULT 13

QY 3 EALGSDYPDPVGAVTLSPLSAVREERKASDGVSGF 39

DB 324 DKIGVSYPVENKVRMYT-ETARMKVALLDDGGVVGF 359

RESULT 15

G01946

nitrinic-oxide synthase (EC 1.14.13.39), neuronal - human

N:Alternate names: nitric oxide synthase 1

C:Species: Homo sapiens (man)

C>Date: 21-Dec-1996 #sequence.revision 06-Jun-1997 #text_change 04-Mar-2000

C:Accession: G01946; I56508; S28878

C:Author: Park, C.; Glanotli, C.; Park, R.; Krishna, G.

S:Submitted to the EMBL Data Library, July 1995

A:Reference number: G08911

A:Accession: G01946

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1433 <PAR>

A:Cross-references: EMBL:U31466; NID:g951318; PIDN:AAB49040.1; PID:g951319

J:Kurashima, Y.; Yokoyama, T.; Yamashita, J.; Esumi, H.

J:Neurochem. 63, 140-145, 1994

A>Title: Expression of two types of nitric oxide synthase mRNA in human neuroblastoma ce

A:Reference number: I56508; MUID:94267447

A:Accession: I56508

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-130, 'K', 132-177, 'LA', 179, 'RP', 182, 'G', 184-1433 <RES>

A:Cross-references: GB:D16408; NID:g506339; PIDN:BAA03895.1; PID:g987662

R:Nakane, M.; Schmidt, H.H.W.; Pollock, J.S.; Foerstermann, U.; Murad, F.

FEBS Lett. 316, 175-180, 1993

A>Title: Cloned human brain nitric oxide synthase is highly expressed in skeletal muscle

A:Reference number: S28878; MUID:93131039

A:Accession: S28878

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-130, 'K', 132-430, 'HR', 493-547, 'L', 549-561, 'A', 563-1362, 'I', 1364-1405, 'I', 14

A:Cross-references: GB:L02861; NID:g189261

C:Genetics:

A:Gene: GDB:NOS1; NOS

A:Cross-references: GDB:I32579; OMIM:163731

A:Map position: 12q24.2-12q24.31

C:Superfamily: GLGF domain homology; flavodoxin homology; NADPH--ferrihemoprotein reduct

C:Keywords: chromoprotein; flavoprotein; heme; iron; metalloprotein; oxidoreductase

F:23-95/Domain: GLGF domain homology <GLG>

F:759-1398/Domain: NADPH--ferrihemoprotein reductase homology <FEH>

F:761-939/Domain: flavodoxin homology <FLX>

F:419/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 22.9%; Score 55.5; DB 2; Length 1433;

Best Local Similarity 47.1%; Pred. No. 48;

Matches 16; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

QY 11 PPVGAVTSLPSAVREERKASDG-SVPFGQLH 43

DB 128 PPTPAVDLSHPAGKEOPLAVDGASGFGNGPOH 161

Search completed: June 1, 2001, 13:46:15
Job time: 57 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2001, 13:45:18 ; Search time 11.63 Seconds
(without alignments)
77.636 Million cell updates/sec

Title: US-09-331-723A-1

Perfect score: 242

Sequence: 1 AAALGSEDFPPVGAATLSY.....ERKASDGVFGQLHPRTQ 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents.AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCUTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142.5	58.9	528	2	US-08-808-931-10
2	142.5	58.9	528	3	US-08-808-323-10
3	142.5	58.9	528	3	US-09-050-603A-10
4	142.5	58.9	528	3	US-09-102-420B-10
5	141.5	58.5	481	1	US-08-472-028A-6
6	141.5	58.5	481	2	US-08-808-931-6
7	141.5	58.5	481	3	US-08-808-323-6
8	141.5	58.5	481	3	US-09-050-603A-6
9	141.5	58.5	481	3	US-09-102-420B-6
10	141.5	58.5	483	4	US-09-071-296-6
11	140.5	58.1	312	2	US-08-808-931-22
12	140.5	58.1	312	3	US-08-808-323-22
13	140.5	58.1	312	3	US-09-050-603A-22
14	140.5	58.1	312	3	US-09-102-420B-22
15	139.5	57.6	539	2	US-08-808-931-16
16	139.5	57.6	539	3	US-08-808-323-16
17	139.5	57.6	539	3	US-09-050-603A-16
18	139.5	57.6	539	3	US-09-102-420B-16
19	138.5	57.2	543	2	US-08-808-931-12
20	138.5	57.2	543	3	US-08-808-323-12
21	138.5	57.2	543	3	US-09-050-603A-12
22	138.5	57.2	543	3	US-09-102-420B-12
23	137.5	56.8	440	2	US-08-808-931-24
24	137.5	56.8	440	3	US-08-808-323-24
25	137.5	56.8	440	3	US-09-050-603A-24
26	137.5	56.8	440	3	US-09-102-420B-24
27	133.5	55.2	537	1	US-08-472-028A-2

28 133.5 55.2 537 2 US-08-808-931-2 Sequence 2, Appli
29 133.5 55.2 537 3 US-08-808-323-2 Sequence 2, Appli
30 133.5 55.2 537 3 US-09-050-603A-2 Sequence 2, Appli
31 133.5 55.2 537 3 US-09-102-420B-2 Sequence 2, Appli
32 133.5 55.2 537 4 US-09-071-296-2 Sequence 2, Appli
33 132.5 54.8 560 2 US-08-808-931-18 Sequence 18, Appli
34 132.5 54.8 560 3 US-08-808-323-18 Sequence 18, Appli
35 132.5 54.8 560 3 US-09-050-603A-18 Sequence 18, Appli
36 132.5 54.8 560 3 US-09-102-420B-18 Sequence 18, Appli
37 131.5 54.3 536 2 US-08-808-931-20 Sequence 20, Appli
38 131.5 54.3 536 3 US-08-808-323-20 Sequence 20, Appli
39 131.5 54.3 536 3 US-09-050-603A-20 Sequence 20, Appli
40 131.5 54.3 536 3 US-09-102-420B-20 Sequence 20, Appli
41 57.5 23.8 1317 3 US-09-083-521-7 Sequence 7, Appli
42 57 23.6 982 2 US-08-673-789-4 Sequence 4, Appli
43 57 23.6 983 1 US-08-162-809-16 Sequence 16, Appli
44 57 23.6 983 1 US-08-167-919A-10 Sequence 10, Appli
45 57 23.6 983 2 US-08-449-645A-21 Sequence 21, Appli

ALIGNMENTS

RESULT 1
US-08-808-931-10
; Sequence 10, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
; TITLE OF INVENTION: Theoreof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5939602artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meligs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 amino acids
; TYPE: amino acid

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-931-6

Query Match 58.5%; Score 141.5; DB 2; Length 481;
Best Local Similarity 61.7%; Pred. No. 3.8e-12;
Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;
QY 1 AAALGSDYPPVAVTSLYPLSAVREERKASDGYPGQLHPTQ 47
DB 297 AADALSFRYPPVAAVTVSPKEAIRKE-CLIDGELQGFGLHPRSQ 342

RESULT 7

US-08-808-323-6
Sequence 6, Application US/08808323
Patent No. 6018105
GENERAL INFORMATION:
APPLICANT: Johnson, Marie
APPLICANT: Volrath, Sandra
APPLICANT: Ward, Eric
TITLE OF INVENTION: Promoters from Plant
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6018105artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 21-JUN-1996
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241

FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-323-6

Query Match 58.5%; Score 141.5; DB 3; Length 481;
Best Local Similarity 61.7%; Pred. No. 3.8e-12;
Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;
QY 1 AAALGSDYPPVAVTSLYPLSAVREERKASDGYPGQLHPTQ 47
DB 297 AADALSFRYPPVAAVTVSPKEAIRKE-CLIDGELQGFGLHPRSQ 342

RESULT 8

US-09-050-603A-6
Sequence 6, Application US/09050603A
Patent No. 6023012
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6023012artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 30-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-050-603A-6

Query Match 58.5%; Score 141.5; DB 3; Length 481;
Best Local Similarity 61.7%; Pred. No. 3.8e-12;
Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 1 AAALGSDYPPVGVAVTSLVPLSAVREERKASDGSVPFGQLHPTQ 47
DB 297 AAALSRFYPPVAAVTVPKPAIRKE-CLIDGELQFGQLHPSQ 342

RESULT 9
US-09-102-420B-6
Sequence 6, Application US/09102420B
Patent No. 6084155
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Heifetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6084155artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC USA
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,420B
FILING DATE: 22-JUN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/059,164
FILING DATE: 13-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,028

FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-102-420B-6

Query Match 58.5%; Score 141.5; DB 3; Length 481;
Best Local Similarity 61.7%; Pred. No. 3.8e-12;
Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 1 AAALGSDYPPVGVAVTSLVPLSAVREERKASDGSVPFGQLHPTQ 47
DB 297 AAALSRFYPPVAAVTVPKPAIRKE-CLIDGELQFGQLHPSQ 342

RESULT 10
US-09-071-296-6
Sequence 6, Application US/09071296
Patent No. 6177245
GENERAL INFORMATION:
APPLICANT: Ward, Eric R
APPLICANT: Volrath, Sandra
TITLE OF INVENTION: Manipulation of Protoporphyrinogen
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,296
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/261,198
FILING DATE: 16-JUN-94
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1748/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-296-6

Query Match 58.5%; Score 141.5; DB 4; Length 483;
Best Local Similarity 61.7%; Pred. No. 3.8e-12;

NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 603012artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,603A
FILING DATE: 30-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-050-603A-22

Query Match 58.1%; Score 140.5; DB 3; Length 312;
Best Local Similarity 61.7%; Pred. No. 3e-12;
Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 1 AAALGSGFDYPPVAVTSLYPLSAVREERKASDGSVPFGQLHPRQ 47
DB 128 ADALSLFYPPVAVTSLYPLSAVREERKASDGSVPFGQLHPRSQ 173

RESULT 14
US-09-102-420B-22
Sequence 22, Application US/09102420B
Patent No. 6084155
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
TITLE OF INVENTION: OXIDASE ("PROTOX")
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6084155artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA

ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,420B
FILING DATE: 22-JUN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/059,164
FILING DATE: 13-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-102-420B-22

Query Match 58.1%; Score 140.5; DB 3; Length 312;
Best Local Similarity 61.7%; Pred. No. 3e-12;
Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 1 AAALGSGFDYPPVAVTSLYPLSAVREERKASDGSVPFGQLHPRQ 47
DB 128 ADALSLFYPPVAVTSLYPLSAVREERKASDGSVPFGQLHPRSQ 173

RESULT 15
US-08-808-931-16
Sequence 16, Application US/08808931
Patent No. 5939602
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Ptoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
TITLE OF INVENTION: Thereof

```

1  NUMBER OF SEQUENCES: 35
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: NO. 5939602artis Corporation
4  STREET: 520 White Plains Road, P.O. Box 2005
5  CITY: Tarrytown
6  STATE: NY
7  COUNTRY: USA
8  * ZIP: 10591-9005
9  COMPUTER READABLE FORM:
10 MEDIUM TYPE: Floppy disk
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: PatentIn Release #1.0, Version #1.30
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/808,931
16 FILING DATE:
17 CLASSIFICATION: 435
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 60/012,705
20 FILING DATE: 28-FEB-1996
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 60/013,612
23 FILING DATE: 28-FEB-1996
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 60/020,003
26 FILING DATE: 21-JUN-1996
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Weigs, J. Timothy
29 REGISTRATION NUMBER: 38,241
30 REFERENCE/DOCKET NUMBER: CGC 1847
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (919) 541-8587
33 TELEFAX: (919) 541-8689
34 INFORMATION FOR SEQ ID NO: 16:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 539 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: not relevant
39 TOPOLOGY: not relevant
40 MOLECULE TYPE: protein
41 PS-08-808-931-16

```

```

Query Match      57.6%; Score 139.5; DB 2; Length 539;
Best Local Similarity 59.6%; Pred. NO. 8.4e-12;
Matches 28; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 1 AAELGSDFYPPVGVATTSYPLSAVREERKASDGVPGFGQLHPRTQ 47
db 355 RADALGSFYPPVAVSTYSYKPAIRRE-CLIDGELKFGQLHPRSQ 400

```

Search completed: June 1, 2001, 13:45:57
Job time: 39 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2001, 13:45:18 ; Search time 18.55 Seconds
(without alignments)
144.834 Million cell updates/sec

Title: US-09-331-723a-1

Perfect score: 242
Sequence: 1 AAEALGSDYPPVGAUTLSY.....ERRKASDGVPGGOLHPRQTQ 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*

```
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	242	100.0	47	W64450	C. reinhardtii por
2	242	100.0	563	Y49538	Chlamydomonas rein
3	242	100.0	563	Y54101	An algal protoporph
4	164.5	68.0	548	Y05200	N. tabacum protopo
5	164.5	68.0	548	Y05201	N. tabacum protopo
6	142.5	58.9	528	W41607	Wheat prototox-1. F
7	142.5	58.9	528	W25738	Wheat protoporphyr
8	141.5	58.5	46	W64452	Z. mays porphyrin
9	141.5	58.5	481	W41605	Maize prototox-1. Z
10	141.5	58.5	481	W25745	Maize protoporphyr
11	141.5	58.5	481	W51255	Maize proto-porphyr

12	141.5	58.5	483	17	R90297
13	140.5 <td>58.1 <td>312</td> <td>18</td> <td>W41612</td> </td>	58.1 <td>312</td> <td>18</td> <td>W41612</td>	312	18	W41612
14	140.5 <td>58.1 <td>312</td> <td>18</td> <td>W25743</td> </td>	58.1 <td>312</td> <td>18</td> <td>W25743</td>	312	18	W25743
15	139.5 <td>57.6 <td>539</td> <td>18</td> <td>W41608</td> </td>	57.6 <td>539</td> <td>18</td> <td>W41608</td>	539	18	W41608
16	139.5 <td>57.6 <td>539</td> <td>18</td> <td>W41609</td> </td>	57.6 <td>539</td> <td>18</td> <td>W41609</td>	539	18	W41609
17	139.5 <td>57.6 <td>539</td> <td>18</td> <td>W25740</td> </td>	57.6 <td>539</td> <td>18</td> <td>W25740</td>	539	18	W25740
18	138.5 <td>57.2 <td>543</td> <td>18</td> <td>W25739</td> </td>	57.2 <td>543</td> <td>18</td> <td>W25739</td>	543	18	W25739
19	138.5 <td>57.2 <td>543</td> <td>20</td> <td>W92498</td> </td>	57.2 <td>543</td> <td>20</td> <td>W92498</td>	543	20	W92498
20	138.5 <td>57.2 <td>543</td> <td>21</td> <td>Y54100</td> </td>	57.2 <td>543</td> <td>21</td> <td>Y54100</td>	543	21	Y54100
21	137.5 <td>56.8</td> <td>440</td> <td>18</td> <td>W41613</td>	56.8	440	18	W41613
22	137.5 <td>56.8</td> <td>440</td> <td>18</td> <td>W25744</td>	56.8	440	18	W25744
23	133.5 <td>55.2</td> <td>46</td> <td>19</td> <td>W64451</td>	55.2	46	19	W64451
24	133.5 <td>55.2</td> <td>413</td> <td>21</td> <td>G42387</td>	55.2	413	21	G42387
25	133.5 <td>55.2</td> <td>416</td> <td>21</td> <td>G42386</td>	55.2	416	21	G42386
26	133.5 <td>55.2</td> <td>537</td> <td>17</td> <td>R90295</td>	55.2	537	17	R90295
27	133.5 <td>55.2</td> <td>537</td> <td>18</td> <td>W41603</td>	55.2	537	18	W41603
28	133.5 <td>55.2</td> <td>537</td> <td>18</td> <td>W25746</td>	55.2	537	18	W25746
29	133.5 <td>55.2</td> <td>537</td> <td>19</td> <td>W51347</td>	55.2	537	19	W51347
30	133.5 <td>55.2</td> <td>537</td> <td>21</td> <td>G42385</td>	55.2	537	21	G42385
31	132.5 <td>54.8</td> <td>560</td> <td>18</td> <td>W41610</td>	54.8	560	18	W41610
32	132.5 <td>54.8</td> <td>560</td> <td>18</td> <td>W25741</td>	54.8	560	18	W25741
33	131.5 <td>54.3</td> <td>536</td> <td>18</td> <td>W41611</td>	54.3	536	18	W41611
34	131.5 <td>54.3</td> <td>536</td> <td>18</td> <td>W25742</td>	54.3	536	18	W25742
35	57	23.6	983	14	R31466
36	57	23.6	983	16	R75711
37	57	23.6	983	21	B08665
38	57	23.6	983	21	B08666
39	56	23.1	346	18	W21730
40	56	23.1	363	18	W23658
41	56	23.1	363	19	W75699
42	56	23.1	363	19	W68185
43	56	23.1	379	21	B58841
44	55.5	22.9	1003	21	B24058
45	55.5	22.9	1004	21	Y66639

ALIGNMENTS

RESULT 1
W64450
ID W64450 standard; Protein; 47 AA.
AC W64450;
DF 16-OCT-1998 (first entry)
XX C. reinhardtii porphyrin herbicide resistance domain protein.
DE
XX Herbicide resistance: porphyrin; protoporphyrinogen oxidase; PPO;
KW crop plant; selective weed control.
XX Chlamydomonas reinhardtii.
OS
XX W09829554-AL.
XX
XX 09-JUL-1998.
XX
XX 27-DEC-1996; 96WO-US20415.
XX
XX 27-DEC-1996; 96WO-US20415.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX (UYDU-) UNIV DUKE.
XX Boynton JE, Gillham NW, Ishige F, Randolph-Anderson BL;
XX Sato R;
XX WPI; 1998-388134/33.
XX DR N-PSDB; V46291.
XX Conferring resistance to herbicides that inhibit proto-porphyrinogen
PT oxidase - by introducing DNA that encodes a resistant form of the

PT enzyme, also resistant plants and algae, mutant DNA, vectors and
 PT transformed microorganisms

PS Claim 1; Page 65; 108pp; English.

XX This sequence represents a fragment of the Chlamydomonas reinhardtii
 CC protoporphyrin herbicide resistance domain which is used in a method to
 CC investigate resistance to herbicides that inhibit protoporphyrinogen
 CC oxidase (PPO). The method is applied to crop plants, allowing use of
 CC herbicides for selective weed control. The method can be used to
 CC identify compounds that are, or are not, PPO-inhibitors.

XX Sequence 47 AA;

Query Match 100.0%; Score 242; DB 19; Length 47;
 Best Local Similarity 100.0%; Pred. No. 8.3e-27;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAALGSDYPPVPGAVTSLVPLSAVREERKASDGSVPFGQLHPTQ 47

Db 1 aaalgsfdyppvgavtlyplsavreerkasdgsvpgfqlhprtq 47

RESULT 2

Y49538
 ID Y49538 standard; Protein: 563 AA.

XX Y49538;

XX 12-JAN-2000 (first entry)

XX Chlamydomonas reinhardtii protoporphyrinogen oxidase.

XX Protoporphyrinogen oxidase; PPO; herbicide; inhibition; resistance;
 KW malignant cell growth; phototropic herbicide.

XX Chlamydomonas reinhardtii.

XX EP955380-A2.

XX 10-NOV-1999.

XX 09-APR-1999; 99EP-0107037.

XX 10-APR-1998; 98JP-0099619.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX Shimokawatoko Y, Nishio S;

XX WPI; 1999-612726/53.

XX N-PSDB; 232119.

XX Evaluating ability of compounds to inhibit protoporphyrinogen oxidase
 PT activity for development of herbicides -

PS Claim 19; Page 26-28; 46pp; English.

XX A method has been developed for evaluating the ability of a compound to
 CC inhibit protoporphyrinogen oxidase (PPO) activity. The method comprises:
 CC (a) culturing a transformant expressing a protoporphyrinogen oxidase
 CC gene in the presence and absence of a test compound to measure a growth
 CC rate; and (b) determining the ability of the compound to inhibit PPO
 CC activity by comparing growth rates. The method is useful for the
 CC development of PPO inhibiting-type herbicides. Also the method is useful
 CC for screening medications having the ability for inhibiting the
 CC malignant cell growth in addition to evaluating the potency of a
 CC phototropic herbicide, screening a compound having a herbicidal
 CC activity. The PPO gene can be used to obtain or generate a gene
 CC resistance to a PPO activity inhibiting agent e.g. a phototropic
 CC herbicide. The present sequence represents Chlamydomonas reinhardtii-
 CC derived PPO.

XX Sequence 563 AA;

Query Match 100.0%; Score 242; DB 20; Length 563;
 Best Local Similarity 100.0%; Pred. No. 1.5e-25;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAALGSDYPPVPGAVTSLVPLSAVREERKASDGSVPFGQLHPTQ 47

Db 377 aaalgsfdyppvgavtlyplsavreerkasdgsvpgfqlhprtq 423

RESULT 3

Y54101

ID Y54101 standard; Protein: 563 AA.

XX Y54101;

XX 27-MAR-2000 (first entry)

XX An algal protoporphyrin IX oxidase protein.

XX Weed control; weed activity; plant; protoporphyrin IX oxidase; PPO;
 KW herbicide; chlormethoxynil; acifluorfen; protoporphyrin X.

XX Chlamydomonas reinhardtii.

XX EP953646-A2.

XX 03-NOV-1999.

XX 30-APR-1999; 99EP-0108463.

XX 30-APR-1998; 98JP-0120553.

XX 02-OCT-1998; 98JP-0281127.

XX 20-NOV-1998; 98JP-0330981.

XX 02-MAR-1999; 99JP-0054730.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX Nakajima H, Nagasawa A;

XX WPI; 2000-088762/08.

XX N-PSDB; 245282.

XX Method of making plants resistant to weed control compounds by
 PT introducing gene encoding protein into plant cell -

XX Example 15; Page 77-79; 119pp; English.

XX The present sequence represents an algal protoporphyrin IX oxidase (PPO).
 CC The sequence was used in the course of the invention. The specification
 CC describes a method of making plants resistant to weed control compounds.
 CC The method comprises introducing a gene encoding a protein into a plant
 CC cell, where the protein has a specific affinity for a substance which
 CC is concerned with the weed activity of a weed control compound,
 CC e.g. protoporphyrin IX; has no ability to modify this protein; and
 CC is virtually free from framework regions in an immunoglobulin; and
 CC expressing the gene. The gene of the invention is used for producing a
 CC plant resistant to specified compounds, such as protoporphyrin IX
 CC oxidase (PPO) inhibitory-type herbicides, e.g. chlormethoxynil,
 CC acifluorfen, etc.. When PPO is inhibited, protoporphyrin IX (the
 CC substrate of PPO) accumulates in the plant cells, and is metabolised to
 CC form protoporphyrin X, followed by formation of active oxygen in the
 CC presence of protoporphyrin X and light, which damages cell functions.

XX Sequence 563 AA;

Query Match 100.0%; Score 242; DB 21; Length 563;

Best Local Similarity 100.0%; Pred. No. 1.5e-25;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX The present sequence is wheat protoporphyrinogen oxidase-1
 CC (protox-1).
 CC The protox-1 promoter can be used to express herbicide resistant
 CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny
 CC containing a chimeric gene of the promoter and a heterologous
 CC coding sequence. The plant can also be used in breeding programmes.
 CC Also hybridising fragments of the protox coding sequence can be
 CC used as probes, e.g. to isolate related genes or for genomic
 CC mapping.
 XX
 XX Sequence 528 AA;
 SQ
 Query Match 58.9%; Score 142.5; DB 18; Length 528;
 Best Local Similarity 61.7%; Pred. No. 1.1e-11;
 Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;
 QY 1 AAALGSFDYPPVGAVTSLYSPLSAVREERKASDGSVPFGQLHPRQTQ 47
 Db 344 aaalskfyppvaavtvsypkeairke-clidgelgfgqlhprsq 389
 RESULT 7
 W25738
 ID W25738 standard; Protein: 528 AA.
 AC W25738;
 XX
 DT 01-MAR-1998 (first entry)
 XX
 DE Wheat protoporphyrinogen oxidase (protox-1).
 XX
 XX Protox-1a; protoporphyrinogen oxidase; inhibitor; wheat;
 KW herbicide tolerance; herbicide resistance; transgenic plant.
 XX
 OS Triticum aestivum cv. Kanzler.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 211 /note= "substitution of Ala-211 by another amino
 FT acid, especially Val or Thr, provides a
 FT modified protox tolerant to a herbicide
 FT (Claims 37-38)"
 FT
 FT Misc-difference 212 /note= "substitution of Gly-212 by another amino
 FT acid, especially Ser, provides a modified
 FT protox tolerant to a herbicide (Claims
 FT 39-40)"
 FT
 FT Misc-difference 356 /note= "substitution of Val-356 by another amino
 FT acid, especially Leu, provides a modified
 FT protox tolerant to a herbicide (Claims
 FT 31-32)"
 FT
 FT Misc-difference 421 /note= "substitution of Ser-421 by another amino
 FT acid, especially Pro, provides a modified
 FT protox tolerant to a herbicide (Claims
 FT 33-34)"
 FT
 FT Misc-difference 466 /note= "substitution of Ile-466 by another amino
 FT acid, especially Thr, provides a modified
 FT protox tolerant to a herbicide (Claims
 FT 41-42)"
 FT
 FT Misc-difference 502 /note= "substitution of Val-502 by another amino
 FT acid, especially Ala, provides a modified
 FT protox tolerant to a herbicide (Claims
 FT 35-36)"
 XX
 PN W09732011-A1.
 XX
 PD 04-SEP-1997.

XX 27-FEB-1997; 97WO-US03313.
 XX
 PR 21-JUN-1996; 96US-0020003.
 PR 28-FEB-1996; 96US-0012705.
 PR 28-FEB-1996; 96US-0013612.
 XX (NOVS) NOVARTIS AG.
 XX
 PI Heifetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;
 XX
 XX WPI; 1997-448683/41.
 DR N-PSDB; T86121.
 XX
 XX New DNA encoding plant protoporphyrinogen oxidase enzyme - and
 PT herbicide resistant mutants, useful to prepare plants resistant to
 PT herbicide which therefore kills undesired vegetation only
 XX
 PS Claim 3; Page 120-123; 196pp; English.
 XX
 CC This protein comprises wheat protoporphyrinogen oxidase (protox-1),
 CC an enzyme that catalyses the oxidation of protoporphyrinogen IX to
 CC protoporphyrin IX. Its amino acid sequence was deduced from a
 CC Protox-1 cDNA clone (see T86121). Sites within the wheat Protox-1
 CC gene have been identified that can be mutated to encode a modified
 CC protox that is resistant to protox inhibitors and hence tolerant of
 CC certain herbicides. Plants, especially crop plants, may be
 CC engineered for resistance to protox inhibitors via mutation of the
 CC native protox gene to a resistant form, or they may be transformed
 CC with a gene encoding an inhibitor-resistant form of a plant protox
 CC enzyme, such as claimed forms from wheat, soybean, cotton,
 CC sugarbeet, oilseed rape, rice and sorghum (see W25738-48).
 CC Application of herbicide will then kill undesired vegetation only.
 CC Protox enzymes can also be expressed in transformed host cells and
 CC used to identify inhibitors of protox enzyme activity, i.e.
 CC herbicide candidates, or to design herbicide tolerant forms of the
 CC enzyme.
 XX
 XX Sequence 528 AA;
 SQ
 Query Match 58.9%; Score 142.5; DB 18; Length 528;
 Best Local Similarity 61.7%; Pred. No. 1.1e-11;
 Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;
 QY 1 AAALGSFDYPPVGAVTSLYSPLSAVREERKASDGSVPFGQLHPRQTQ 47
 Db 344 aaalskfyppvaavtvsypkeairke-clidgelgfgqlhprsq 389
 RESULT 8
 W64452
 ID W64452 standard; Protein: 46 AA.
 XX
 AC W64452;
 XX
 DT 16-OCT-1998 (first entry)
 XX
 DE Z. mays porphyrin herbicide resistance domain protein.
 XX
 KW Herbicide resistance; porphyrin; protoporphyrinogen oxidase; PPO;
 KW crop plant; selective weed control; corn.
 XX
 OS Zea mays.
 XX
 PN W09829554-A1.
 XX
 PD 09-JUL-1998.
 XX
 PF 27-DEC-1996; 96WO-US20415.
 XX
 PR 27-DEC-1996; 96WO-US20415.
 XX

(SUMO) SUMITOMO CHEM CO LTD.
(UTDU-) UNIV DUKE.

PA
XX PA
XX PI
PI Boynton JE, Gillham NW, Ishige F, Randolph-Anderson BL;
XX Sato R;
XX WPI; 1998-388134/33.
DR N-PSDB; V46293.
XX

Conferring resistance to herbicides that inhibit proto-porphyrinogen oxidase - by introducing DNA that encodes a resistant form of the enzyme, also resistant plants and algae, mutant DNA, vectors and transformed microorganisms

XX
XX Claim 1; Page 66; 108pp; English.
PS
XX This sequence represents a fragment of the Zea mays strain B73 inbred corn porphyrin herbicide resistance domain which is used in a method to investigate resistance to herbicides that inhibit protoporphyrinogen oxidase (PPO). The method is applied to crop plants, allowing use of herbicides for selective weed control. The method can be used to identify compounds that are, or are not, PPO-inhibitors.

XX Sequence 46 AA;

Query Match 58.5%; Score 141.5; DB 19; Length 46;
Best Local Similarity 61.7%; Pred. No. 9.2e-13;
Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps

QY 1 AAEALGSDFFPPVPGAVTILSYPLSAVRERKASDGSVPFGQLHPRTQ 47
||:|| ||| ||| ||| ||| :||: |||||||||:
Db 1 aaalsrfyppvaavtsypkairke-clldgeiqfgqlhpsrq 46

RESULT 9

ID W41605
XX W41605 standard; Protein; 481 AA.
XX W41605;

DT 20-APR-1998 (first entry)
DE Maize protox-1.

KW Protoporphyrinogen oxidase-1; protox-1; promoter; maize;
KW herbicide resistance; breeding programme; probe; gene isolation;
KW genomic mapping.

OS Zea mays.
XX WO9732028-A1.
XX 04-SEP-1997.

PF 27-FEB-1997; 97WO-US03343.
PR 21-JUN-1996; 96US-0020003.
PR 28-FEB-1996; 96US-0012705.
PR 28-FEB-1996; 96US-0013612.

XX (NOVS) NOVARTIS AG.
XX Johnson MA, Volrath SL, Ward ER;
XX WPI; 1997-489209/45.
DR N-PSDB; V04306.
XX

DNA containing a plant proto-porphyrinogen oxidase gene promoter - optionally linked to a heterologous gene, especially to express herbicide-resistant enzymes, and plants containing such constructs

XX Claim 29; Pages 49-52; 114pp; English.
PS

XX	The present sequence is maize protoporphyrinogen oxidase-1
CC	(protox-1).
CC	The protox-1 promoter can be used to express herbicide resistant
CC	enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC	containing a chimeric gene of the promoter and a heterologous
CC	coding sequence. The plant can also be used in breeding programmes.
CC	Also hybridising fragments of the protox coding sequence can be
CC	used as probes, e.g. to isolate related genes or for genomic
CC	mapping.
XX	
XX	Sequence 481 AA;
SQ	
	Query Match 58.5%; Score 141.5; DB 18; Length 481;
	Best Local Similarity 61.7%; Pred. No. 1.4e-11;
	Matches 29; Conservative 6; Mismatches 11; Indels 1; Gap
QY	1 AAEALGSDFPYPGVAVTSLPSVAREERKASDGPVGFGQLHPRTO 47
	:
DB	297 aaadalsrfypypvaavtvsypkeairke-clidgelqgfgqlhprsq 342
RESULT 10	
W25745	ID W25745 standard; Protein; 481 AA.
ID	AC AC
XX	W25745:
XX	01-MAR-1998 (first entry)
XX	
DE	Maize protoporphyrinogen oxidase (protox-1).
DE	
KW	Protox-1; protoporphyrinogen oxidase; inhibitor; maize;
KW	herbicide tolerance; herbicide resistance; transgenic plant.
XX	
XX	Zea mays.
OS	
FH	Key Location/Qualifiers
FT	Misc-difference 159 /note= "Substitution of Cys-159 by another amino
FT	acid, especially Phe or Lys, provides a
FT	modified protox tolerant to a herbicide
FT	(Claims 19-20)"
FT	Misc-difference 164 /note= "Substitution of Ala-164 by another amino
FT	acid, especially Thr, Leu or Val, provides
FT	a modified protox tolerant to a herbicide
FT	(Claims 25-26)"
FT	Misc-difference 165 /note= "Substitution of Gly-165 by another amino
FT	acid, especially Ser or Leu, provides a
FT	modified protox tolerant to a herbicide
FT	(Claims 27-28)"
FT	Misc-difference 370 /note= "Substitution of Tyr-370 by another amino
FT	acid, especially Ile or Met, provides a
FT	modified protox tolerant to a herbicide
FT	(Claims 29-30)"
FT	Misc-difference 419 /note= "Substitution of Ile-419 by another amino
FT	acid, especially Thr, His, Gly or Asn
FT	provides a modified protox tolerant to a
FT	herbicide (Claims 22-23)"
PN	
XX	WO9732011-A1.
XX	
PD	04-SEP-1997.
XX	
PF	27-FEB-1997; 97WO-US03313.
XX	
PR	21-JUN-1996; 96US-0020003.
PR	28-FEB-1996; 96US-0012705.

CC DNA encoding a modified protoporphyrinogen oxidase (PPO) or a
 CC chimeric gene comprising a promoter (pref. active in a plant) linked
 CC to such a DNA or DNA encoding a wild type PPO can be used (1) to
 CC impart herbicide resistance to plants; (2) for treating and
 CC diagnosing deficient PPO activity in animals (esp. variegate
 CC porphyria); and (3) for the production of recombinant PPO which is
 CC useful as an assay reagent and in rational design of new inhibitory
 CC herbicides. Herbicide resistant PPO genes can also be used to
 CC select plants transformed with a transgene and probes derived from
 CC the genes can be used to quantify levels of PPO mRNA. This partial
 CC sequence of maize protox-1 does have PPO activity.
 XX
 SQ Sequence 483 AA;

Query Match 58.5%; Score 141.5; DB 17; Length 483;
 Best Local Similarity 61.7%; Pred. No. 1.4e-11;
 Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;
 QY 1 AAEALGSDYPPVPGAVTSLSYPLSAVREERKASDGSVPFGQLHPRTQ 47
 ||:| | ||| ||| ||| ||:| | | : ||||| |||
 Db 299 aadaisfyyppvaavtvpkkaikr-clldgelqgfgqlhprsq 344

RESULT 13
 W41612
 ID W41612 standard; Protein; 312 AA.
 AC W41612;
 XX
 XX
 DT 20-APR-1998 (first entry)
 XX
 DE Rice protox-1.
 XX
 KW Protoporphyrinogen oxidase-1; protox-1; promoter; rice;
 KW herbicide resistance; breeding programme; probe; gene isolation;
 KW genomic mapping.
 XX
 OS Oryza sativa.
 XX
 FN WO9732028-Al.
 XX
 PD 04-SEP-1997.

XX 27-FEB-1997; 97WO-US03343.
 XX 21-JUN-1996; 96US-0020003.
 PR 28-FEB-1996; 96US-0012705.
 PR 28-FEB-1996; 96US-0013612.
 XX
 PA (NOVS) NOVARTIS AG.
 XX
 XX Johnson MA, Volrath SL, Ward ER;
 XX WPI: 1997-489209/45.
 DR N-PSDB; V04316.
 DR
 XX
 XX DNA containing a plant proto-porphyrinogen oxidase gene promoter -
 PT optionally linked to a heterologous gene, especially to express
 PT herbicide-resistant enzymes, and plants containing such constructs
 XX
 PS Claim 41; Pages 93-95; 114pp; English.
 XX
 CC The present sequence is rice protoporphyrinogen oxidase-1
 CC (protax-1).
 CC The protax-1 promoter can be used to express herbicide resistant
 CC enzymes, specifically protax, i.e. a plant tissue, plant or progeny
 CC containing a chimeric gene of the promoter and a heterologous
 CC coding sequence. The plant can also be used in breeding programmes.
 CC Also hybridising fragments of the protax coding sequence can be
 CC used as probes, e.g. to isolate related genes or for genomic
 CC mapping.

SQ Sequence 312 AA;

Query Match 58.1%; Score 140.5; DB 18; Length 312;
 Best Local Similarity 61.7%; Pred. No. 1.2e-11;
 Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;
 QY 1 AAEALGSDYPPVPGAVTSLSYPLSAVREERKASDGSVPFGQLHPRTQ 47
 ||:| | ||| ||| ||| ||:| | | : ||||| |||
 Db 128 aadaisfyyppvaavtvpkkaikr-clldgelqgfgqlhprsq 173

RESULT 14
 W25743
 ID W25743 standard; Protein; 312 AA.
 AC W25743;
 XX
 XX
 DT 01-MAR-1998 (first entry)
 XX
 DE Rice protoporphyrinogen oxidase (protax-1).
 XX
 KW Protax-1; protoporphyrinogen oxidase; inhibitor; rice;
 KW herbicide tolerance; herbicide resistance; transgenic plant.
 XX
 OS Oryza sativa.
 XX
 FN WO9732011-Al.
 XX
 PD 04-SEP-1997.

XX 27-FEB-1997; 97WO-US03313.

XX 21-JUN-1996; 96US-0020003.
 PR 28-FEB-1996; 96US-0012705.
 PR 28-FEB-1996; 96US-0013612.

XX (NOVS) NOVARTIS AG.

XX Heifetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;
 DR WPI: 1997-448693/41.
 DR N-PSDB; T86126.

XX New DNA encoding plant protoporphyrinogen oxidase enzyme - and
 PT herbicide resistant mutants, useful to prepare plants resistant to
 PT herbicide which therefore kills undesired vegetation only

PS Claim 13; Page 152-154; 196pp; English.

XX This protein comprises rice protoporphyrinogen oxidase
 CC (protax-1), an enzyme that catalyses the oxidation of
 CC protoporphyrinogen IX to protoporphyrin IX. Its amino acid
 CC sequence was deduced from an isolated protax-1 cDNA clone (see
 CC T86126). The rice protax-1 can be modified to render it
 CC resistant to protax inhibitors and hence tolerant of certain
 CC herbicides. Plants, especially crop plants, may be engineered for
 CC resistance to protax inhibitors via mutation of the native protax
 CC gene to a resistant form, or they may be transformed with a gene
 CC encoding an inhibitor-resistant form of a plant protax enzyme, such
 CC as claimed forms from wheat, soybean, cotton, sugarbeet, oilseed
 CC rape, rice and sorghum (see W25738-48). Application of herbicide
 CC will then kill undesired vegetation only. Protax enzymes can also
 CC be expressed in transformed host cells and used to identify
 CC inhibitors of protax enzyme activity, i.e. herbicide candidates, or
 CC to design herbicide tolerant forms of the enzyme.

SQ Sequence 312 AA;

Query Match 58.1%; Score 140.5; DB 18; Length 312;
 Best Local Similarity 61.7%; Pred. No. 1.2e-11;
 Matches 29; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 1, 2001, 23:57:42 ; Search time 1131.07 Seconds
(without alignments)
1089.048 Million cell updates/sec

Title: US-09-331-723A-4
Perfect score: 141
Sequence: 1 gccgcgagccctgggtc.....agctgacccgcgcgcgacgcag 141

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est25.*
26: gb_est26.*
27: gb_est27.*
28: gb_est28.*
29: gb_est29.*
30: gb_est30.*
31: gb_est31.*
32: gb_est32.*
33: gb_est33.*
34: gb_est34.*
35: gb_est35.*
36: gb_est36.*
37: gb_est37.*
38: gb_est38.*
39: gb_est39.*
40: gb_est40.*
41: gb_est41.*
42: gb_est42.*
43: gb_est43.*
44: gb_est44.*
45: gb_est45.*
46: gb_est46.*
47: gb_est47.*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: em_gss_pln1:*
191: em_gss_pln2:*
192: em_gss_pro:*
193: em_gss_rod1:*
194: em_gss_rod2:*
195: em_gss_rod3:*
196: em_gss_rod4:*
197: em_gss_rod5:*
198: em_gss_vit1:*
199: em_gss_vit2:*
200: em_gss_vit3:*
201: gb_gss1:*
202: gb_gss2:*
203: gb_gss3:*
204: gb_gss4:*
205: gb_gss5:*
206: gb_gss6:*
207: gb_gss7:*
208: gb_gss8:*
209: gb_gss9:*
210: gb_gss10:*
211: gb_gss11:*
212: gb_gss12:*
213: gb_gss13:*
214: gb_gss14:*
215: gb_gss15:*
216: gb_gss16:*
217: gb_gss17:*
218: gb_gss18:*
219: gb_gss19:*
220: gb_gss20:*
221: gb_gss21:*
222: gb_gss22:*
223: gb_gss23:*
224: gb_gss24:*
225: gb_gss25:*
226: gb_gss26:*
227: gb_gss27:*
228: gb_gss28:*
229: gb_gss29:*
230: gb_gss30:*
231: gb_gss31:*
232: gb_gss32:*
233: gb_gss33:*
234: gb_gss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %			DB ID	Description
		Match	Length			
1	141	100.0	533	163	BE129089	BE129089 894020B03
2	43	30.5	500	113	AW256120	AW256120 687062E03
3	42.4	30.1	932	229	CNS00720	AL066742 Drosophil
4	41.4	29.4	466	117	AW566109	AW566109 660062B03
5	41.4	29.4	577	116	AW498392	AW498392 660047C10
6	39.8	28.2	508	104	AJ280717	AJ280717 4A3A-AAS-
7	39.8	28.2	1375	119	AW727483	AW727483 GA_Ea001
8	38.6	27.4	651	105	AL508953	AL508953 AL508953
9	37.6	26.7	430	102	AI815951	AI815951 au44f02.y
10	37.4	26.5	179	9	AA594531	AA594531 nl94h11.s
11	37.2	26.4	401	175	D22780	D22780 RICCI302A.R
12	37.2	26.4	925	229	CNS0091P	AL053013 Drosophil
13	37	26.2	464	31	AV620164	AV620164 AV620164
14	36.8	26.1	846	231	CNS04J01	AL293770 Tetraodon
15	36.8	26.1	945	230	CNS02D8L	AL192126 Tetraodon
16	36.6	26.0	513	105	AL450546	AL450546 AL450546
17	36.6	26.0	925	229	CNS0091P	AL053013 Drosophil
18	36.6	26.0	1039	230	CNS02V8F	AL215448 Tetraodon


```

DEFINITION 660047C10.xl 660 - Mixed stages of anther and pollen Zea mays cDNA,
            mRNA sequence.
ACCESSION  AW498392
VERSION     AW498392.1 GI:7136035
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae
            ; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 577)
AUTHORS    Walbot,V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL    University
COMMENT    Unpublished (1999)
            Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 660047 row: C column: 10.
FEATURES    Location/Qualifiers
             1..577
             /organism="Zea mays"
             /cultivar="Ohio43"
             /db_xref="taxon:4577"
             /clone_lib="660 - Mixed stages of anther and pollen"
             /tissue_type="whole premeiotic anthers to pollen shed"
             /dev_stage="premeiotic anthers to pollen shed"
             /lab_host="XL0LR"
             /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
             Site_2: XhoI; Anther and pollen cDNA library.
             Directionally sequenced with 5' end at the EcoRI site.
             Created by Amie Franklin."
BASE COUNT  82 a 180 c 199 g 116 t
ORIGIN
Query Match      29.4%; Score 41.4; DB 116; Length 577;
Best Local Similarity 56.1%; Pred. No. 1.9;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy  2  ccgcgagccgtggtccttcgactaccgcgcgtgagcgccgacgctcgtcgtacc 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  420  CCcCGGGTGTGGGCGCCAGGCAAGCCGTGCGGTGCGCACCGGTGAGGCTGTAGTCCT 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  62  cgctgagcgccgtgcgggagagcgcaagcctcgagcgggtccggtcgcgggcttcgtgc 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  480  CGCGCGCGGCTTGCAGAGACGCGCACGCGGTGCGGGCTCCCGCGTGGCGCGGAGG 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  122  agctgcacccgcgcgcga 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  540  AGTTGGCTCGGGACCCCA 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
AJ280717
LOCUS      AJ280717 508 bp mRNA EST 30-JUN-2000
DEFINITION 4A3A-AAS-G-09-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-AAS-G-09, mRNA sequence.
ACCESSION  AJ280717
VERSION     AJ280717.1 GI:6928598
KEYWORDS    EST.
SOURCE      African malaria mosquito.
ORGANISM    Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1 (bases 1 to 508)
AUTHORS    Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansong,W., Soares,M.B.

```

```

and Kafatos,F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES    Location/Qualifiers
             1..508
             /organism="Anopheles gambiae"
             /strain="4A r/r"
             /db_xref="taxon:7165"
             /clone_lib="4A3A-AAS-G-09"
             /cell_line="Anopheles gambiae immune competent 4A3A"
             /lab_host="E. coli DH10B"
             /note="Vector: p7T3D-Pac (Pharmacia) with a modified
             polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
             forward priming site which reads from the 3' end of the
             cDNA. The 4A3A is a directionally cloned and normalized
             cDNA library that was constructed from the 4A3A cell line
             oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
             (1996) : Normalization and Subtraction: Two approaches to
             Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT  58 a 168 c 175 g 107 t
ORIGIN
Query Match      28.2%; Score 39.8; DB 104; Length 508;
Best Local Similarity 64.8%; Pred. No. 4.5;
Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy  46  gtgacgtctgtaccgcgtgagcgcgtgcgggagcgcaagcctcgacgggtcc 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  146  GAGTCGAAGTCCTCCCGCGGACAGGAGGAGGAGGTGCGGAACACCTCGTACGTGTCG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  106  gtgcgggcttcggtcagctgcacccgcgcga 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  206  GTGCGGCGACGCGGAGGAGCTCGCGCGGGGCA 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AW727483/c
LOCUS      AW727483 1375 bp mRNA EST 16-NOV-2000
DEFINITION GA_Ea0012118 Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA_Ea0012118, mRNA sequence.
ACCESSION  AW727483
VERSION     AW727483.1 GI:7625038
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
ORGANISM    Gossypium arboreum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Malvales; Malvaceae; Gossypium.
REFERENCE   1 (bases 1 to 1375)
AUTHORS    Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE      An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
JOURNAL    Unpublished (2000)
COMMENT    Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            High quality sequence stop: 1375.
            Location/Qualifiers
             1..1375

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2001, 00:04:52 ; Search time 128.09 Seconds
(without alignments)
642.620 Million cell updates/sec

Title: US-09-331-723A-4

Perfect score: 141
Sequence: 1 gccgcagccctggctc.....agctgcacccgcgcacgag 141

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 678276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/NA2001.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	141	100.0	141	19	V46291
2	141	100.0	1838	20	232119
3	141	100.0	1838	21	245282
4	139.4	98.9	2573	19	V46297
5	139.4	98.9	3381	19	V46298
6	136.2	96.6	3381	18	T78597
7	136.2	96.6	3383	18	T60965
8	38.8	27.5	2900	13	Q20372
9	38.2	27.1	30001	18	T61016
10	38.2	27.1	30001	20	X05110
11	38	27.0	2310	20	221502

C	12	35.8	25.4	1288	20	V84525
	13	35.4	25.1	3035	21	A63462
	14	35.4	25.1	42000	21	A63349
	15	35.4	25.1	63164	21	A63348
C	16	34.6	24.5	1257	21	A75110
	17	34.6	24.5	1286	21	A26456
C	18	34.6	24.5	1342	21	A26424
	19	34.2	24.3	2657	20	X84348
	20	34.2	24.3	3510	21	C65299
	21	34.2	24.3	43280	18	T80413
C	22	34	24.1	1215	18	T73246
	23	34	24.1	1329	21	A4984
C	24	34	24.1	2855	20	X24071
	25	33.8	24.0	1874	20	X28279
	26	33.8	24.0	1874	20	X28278
	27	33.8	24.0	8937	13	Q20502
	28	33.8	24.0	10706	18	T46941
C	29	33.6	23.8	24379	18	T93095
	30	33.6	23.8	24379	19	V25925
	31	33.4	23.7	738	20	X86968
	32	33.4	23.7	2490	18	T91358
	33	33.4	23.7	2634	19	V23480
	34	33.4	23.7	2634	21	A13906
	35	33.4	23.7	6436	20	X86965
	36	33.4	23.7	17612	19	V23494
	37	33.4	23.7	17612	21	A13905
	38	33.4	23.7	58857	21	A58471
	39	33.2	23.5	780	18	T76882
	40	33.2	23.5	900	18	T76877
C	41	33.2	23.5	1215	19	V23483
C	42	33.2	23.5	1215	21	A13894
	43	33.2	23.5	5516	18	T76880
C	44	33.2	23.5	17612	19	V23494
	45	33.2	23.5	17612	21	A13905

ALIGNMENTS

RESULT 1
V46291
ID V46291 standard; DNA; 141 BP.
XX
AC V46291;
XX
DT 16-OCT-1998 (first entry)
XX
DE C. reinhardtii porphyrin herbicide resistance domain genomic DNA.
XX
KW Herbicide resistance; porphyrin; protoporphyrinogen oxidase; PPO;
XX crop plant; selective weed control; ss.
XX
OS Chlamydomonas reinhardtii.
XX
PN W09829554-A1.
XX
PD 09-JUL-1998.
XX
PF 27-DEC-1996; 96WO-US20415.
XX
PR 27-DEC-1996; 96WO-US20415.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX (UYDU-) UNIV DUKE.
XX Boynton JE, Gillham NW, Ishige F, Randolph-Anderson BL;
XX Sato R;
XX WPI: 1998-388134/33.
XX P-PSDB; W64450.
XX
XX Conferring resistance to herbicides that inhibit proto-porphyrinogen
XX oxidase - by introducing DNA that encodes a resistant form of the

PT enzyme, also resistant plants and algae, mutant DNA, vectors and
PT transformed microorganisms
PS Example 10; Page 67; 108pp; English.
XX
XX This genomic DNA sequence encodes a fragment of the Chlamydomonas
-reinhardtii porphyrin herbicide resistance domain which is used in a
CC method to investigate resistance to herbicides that inhibit
CC protoporphyrinogen oxidase (PPO). The method is applied to crop plants,
CC allowing use of herbicides for selective weed control. The method can be
CC used to identify compounds that are, or are not, PPO-inhibitors.
XX
SQ Sequence 141 BP; 15 A; 53 C; 54 G; 19 T; 0 other;

Query Match 100.0%; Score 141; DB 19; Length 141;
Best Local Similarity 100.0%; Pred. No. 2e-22;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 gccgcagagccctggctcttcgactaccgcgcggtggcgccgtgacgtgtctgtac 60
|||||
Db 1 gccgcagagccctggctcttcgactaccgcgcggtggcgccgtgacgtgtctgtac 60
|||||
Qy 61 ccgctgagcgcgtgcggagagcgcaaggcctcgagcgggtccgtgccgggcttcggt 120
|||||
Db 61 ccgctgagcgcgtgcggagagcgcaaggcctcgagcgggtccgtgccgggcttcggt 120
|||||
Qy 121 cagctgcacccgcgcacgcag 141
|||||
Db 121 cagctgcacccgcgcacgcag 141
|||||

RESULT 2

232119
ID 232119 standard; cDNA; 1838 BP.
XX
AC 232119;
XX
DT 12-JAN-2000 (first entry)
XX
DE Chlamydomonas reinhardtii protoporphyrinogen oxidase encoding cDNA.
XX
XX Protoporphyrinogen oxidase; PPO; herbicide; inhibition; resistance;
KW malignant cell growth; phototropic herbicide; ss.
XX
OS Chlamydomonas reinhardtii.
XX
FH Key Location/Qualifiers
FT CDS 2..1793
FT /*tag= a
FT /product= "protoporphyrinogen oxidase"
XX
PN EP955380-A2.
XX
PD 10-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107037.
XX
PR 10-APR-1998; 98JP-0099619.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Shimokawato Y, Nishio S;
XX
XX WPI; 1999-612726/53.
DR P-PSDB; Y49538.
XX
XX Evaluating ability of compounds to inhibit protoporphyrinogen oxidase
PT activity for development of herbicides -
XX
XX Claim 22; Page 28-31; 46pp; English.
XX
XX A method has been developed for evaluating the ability of a compound to

CC inhibit protoporphyrinogen oxidase (PPO) activity. The method comprises:
CC (a) culturing a transformant expressing a protoporphyrinogen oxidase
CC gene in the presence and absence of a test compound to measure a growth
CC rate; and (b) determining the ability of the compound to inhibit PPO
CC activity by comparing growth rates. The method is useful for the
CC development of PPO inhibiting-type herbicides. Also the method is useful
CC for screening medications having the ability for inhibiting the
CC malignant cell growth in addition to evaluating the potency of a
CC phototropic herbicide, screening a compound having a herbicidal
CC activity. The PPO gene can be used to obtain or generate a gene
CC resistance to a PPO activity inhibiting agent e.g. a phototropic
CC herbicide. The present sequence encodes Chlamydomonas reinhardtii-
XX derived PPO.
XX
SQ Sequence 1838 BP; 302 A; 619 C; 634 G; 283 T; 0 other;

Query Match 100.0%; Score 141; DB 20; Length 1838;
Best Local Similarity 100.0%; Pred. No. 1.6e-22;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 gccgcagagccctggctcttcgactaccgcgcggtggcgccgtgacgtgtctgtac 60
|||||
Db 1130 gccgcagagccctggctcttcgactaccgcgcggtggcgccgtgacgtgtctgtac 1189
|||||
Qy 61 ccgctgagcgcgtgcggagagcgcaaggcctcgagcgggtccgtgccgggcttcggt 120
|||||
Db 1190 ccgctgagcgcgtgcggagagcgcaaggcctcgagcgggtccgtgccgggcttcggt 1249
|||||
Qy 121 cagctgcacccgcgcacgcag 141
|||||
Db 1250 cagctgcacccgcgcacgcag 1270
|||||

RESULT 3

245282
ID 245282 standard; DNA; 1838 BP.
XX
AC 245282;
XX
DT 27-MAR-2000 (first entry)
XX
DE DNA encoding an algal protoporphyrin IX oxidase protein.
XX
XX Weed control; weed activity; plant; protoporphyrin IX oxidase; PPO;
KW herbicide; chlormethoxynil; acifluorfen; protoporphyrin X; ss.
XX
OS Chlamydomonas reinhardtii.
XX
FH Key Location/Qualifiers
FT CDS 2..1693
FT /*tag= a
FT /product= "protoporphyrin IX oxidase"
XX
PN EP953646-A2.
XX
PD 03-NOV-1999.
XX
PF 30-APR-1999; 99EP-0108463.
XX
PR 30-APR-1998; 98JP-0120553.
PR 02-OCT-1998; 98JP-0281127.
PR 20-NOV-1998; 98JP-0330981.
PR 02-MAR-1999; 99JP-0054730.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nakajima H, Nagasawa A;
XX
XX WPI; 2000-088762/08.
DR P-PSDB; Y54101.
XX
XX Method of making plants resistant to weed control compounds by

PT introducing gene encoding protein into plant cell -
 XX Example 15; Page 72-76; 119pp; English.

XX The present sequence encodes an algal protoporphyrin IX oxidase (PPO).
 CC The sequence was used in the course of the invention. The specification
 CC describes a method of making plants resistant to weed control compounds.
 CC The method comprises introducing a gene encoding a protein into a plant
 CC cell, where the protein has a specific affinity for a substance which
 CC is concerned with the weed activity of a weed control compound, and
 CC e.g. protoporphyrin IX; has no ability to modify this protein; and
 CC is virtually free from framework regions in an immunoglobulin; and
 CC expressing the gene. The gene of the invention is used for producing a
 CC plant resistant to specified compounds, such as protoporphyrin IX
 CC oxidase (PPO) inhibitory-type herbicides, e.g. chlormethoxynil,
 CC acifluorfen, etc.. When PPO is inhibited, protoporphyrin IX (the
 CC substrate of PPO) accumulates in the plant cells, and is metabolized to
 CC form protoporphyrin X, followed by formation of active oxygen in the
 CC presence of protoporphyrin X and light, which damages cell functions.

XX Sequence 1838 BP; 302 A; 619 C; 534 G; 283 T; 0 other;

Query Match 100.0%; Score 141; DB 21; Length 1838;
 Best Local Similarity 100.0%; Pred. No. 1.6e-22;
 Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gccgccgagccctgggtccttcgactaccgccggtggcgccgtgacgtctgtac 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1130 gccgccgagccctgggtccttcgactaccgccggtggcgccgtgacgtctgtac 1189
 QY 61 ccgctgagcgcctgcggagagcgcaagcctcgagcgggtccggtccgggttcggt 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1190 ccgctgagcgcctgcggagagcgcaagcctcgagcgggtccggtccgggttcggt 1249
 QY 121 cagctgacccgcgcacgacg 141
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1250 cagctgacccgcgcacgacg 1270

RESULT 4

ID V46297 standard; DNA; 2573 BP.
 AC V46297;
 XX V46297;
 DT 16-OCT-1998 (first entry)
 XX C. reinhardtii protoporphyrinogen oxidase DNA fragment Xho/Pmac2.6.
 DE
 XX Herbicide resistance; porphyrin; protoporphyrinogen oxidase; PPO;
 KW crop plant; selective weed control; ss.
 XX Chlamydomonas reinhardtii.
 OS
 PN W09829554-A1.
 XX
 PD 09-JUL-1998.
 XX
 PF 27-DEC-1996; 96WO-US20415.
 XX
 PR 27-DEC-1996; 96WO-US20415.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX (UYDU-) UNIV DUKE.
 PI Boynton JE, Gillham NW, Ishige F, Randolph-Anderson BL;
 XX -Sato R;
 DR WPI; 1998-388134/33.
 XX

XX Conferring resistance to herbicides that inhibit proto-porphyrinogen
 PT oxidase - by introducing DNA that encodes a resistant form of the

PT enzyme, also resistant plants and algae, mutant DNA, vectors and
 PT transformed microorganisms

PS Example 7; Page 70-71; 108pp; English.

XX This genomic DNA sequence encodes protoporphyrinogen oxidase from
 CC Chlamydomonas reinhardtii (fragment Xho/Pmac2.6). This sequence is used
 CC in a method which investigates resistance to herbicides that inhibit
 CC protoporphyrinogen oxidase (PPO) using a porphyrin herbicide resistance
 CC domain. The method is applied to crop plants, allowing use of herbicides
 CC for selective weed control. The method can be used to identify compounds
 CC that are, or are not, PPO-inhibitors.

XX Sequence 2573 BP; 470 A; 859 C; 735 G; 509 T; 0 other;

Query Match 98.9%; Score 139.4; DB 19; Length 2573;
 Best Local Similarity 99.3%; Pred. No. 3.4e-22;
 Matches 140; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gccgccgagccctgggtccttcgactaccgccggtggcgccgtgacgtctgtac 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 255 gccgccgagccctgggtccttcgactaccgccggtggcgccgtgacgtctgtac 314
 QY 61 ccgctgagcgcctgcggagagcgcaagcctcgagcgggtccggtccgggttcggt 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 315 ccgctgagcgcctgcggagagcgcaagcctcgagcgggtccggtccgggttcggt 374
 QY 121 cagctgacccgcgcacgacg 141
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 375 cagctgacccgcgcacgacg 395

RESULT 5

ID V46298 standard; DNA; 3381 BP.
 AC V46298;
 XX V46298;
 DT 16-OCT-1998 (first entry)
 XX C. reinhardtii protoporphyrinogen oxidase DNA fragment Xho3.4.
 DE
 XX Herbicide resistance; porphyrin; protoporphyrinogen oxidase; PPO;
 KW crop plant; selective weed control; ss.
 XX Chlamydomonas reinhardtii.
 OS
 PN W09829554-A1.
 XX
 PD 09-JUL-1998.
 XX
 PF 27-DEC-1996; 96WO-US20415.
 XX
 PR 27-DEC-1996; 96WO-US20415.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX (UYDU-) UNIV DUKE.
 PI Boynton JE, Gillham NW, Ishige F, Randolph-Anderson BL;
 XX Sato R;
 DR WPI; 1998-388134/33.
 XX

XX Conferring resistance to herbicides that inhibit proto-porphyrinogen

PT oxidase - by introducing DNA that encodes a resistant form of the
 PT enzyme, also resistant plants and algae, mutant DNA, vectors and
 PT transformed microorganisms

CC in a method which investigates resistance to herbicides that inhibit
CC protoporphyrinogen oxidase (PPO) using a porphyrin herbicide resistance
CC domain. The method is applied to crop plants, allowing use of herbicides
CC for selective weed control. The method can be used to identify compounds
CC that are, or are not, PPO-inhibitors.

XX SQ Sequence 3381 BP; 544 A; 1058 C; 1010 G; 569 T; 0 other;

Query Match 98.9%; Score 139.4; DB 19; Length 3381;
Best Local Similarity 99.3%; Pred. No. 3.3e-22;
Matches 140; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gccgcgagccctgggctcttcgactaccgcgagcggtggcgccgctgacgtgctgac 60
|||||
Db 255 gccgcgagccctgggctcttcgactaccgcgagcggtggcgccgctgacgtgctgac 314
|||||
Qy 61 ccgctgagcgcgtgagcgagcgcaaggcctcgacgagcggtgctgcccggctcggt 120
|||||
Db 315 ccgctgagcgcgtgagcgagcgcaaggcctcgacgagcggtgctgcccggctcggt 374
|||||
Qy 121 cagctgacccgcgcacgacg 141
|||||
Db 375 cagctgacccgcgcacgacg 395
|||||

RESULT 6

T78597/c
ID T78597 standard; DNA; 3381 BP.

XX AC T78597;

XX DT 03-NOV-1997 (first entry)

XX DE Porphyrin-accumulating type herbicide resistant DNA fragment.

XX KW Plant; algal; transformation; porphyrin herbicide; weed control;
XX KW oxadiazon; oxyfluorfen; ss.

XX OS Chlamydomonas reinhardtii - strain RS-3.

XX PN W09704089-A2.

XX PD 06-FEB-1997.

XX PF 19-JUL-1996; 96WO-US11999.

XX PR 20-JUL-1995; 95WO-US09098.

XX XX (SUMO) SUMITOMO CHEM CO LTD.
XX PA (UYDU-) UNIV DUKE.

XX PI Boynton J, Gillham NW, Harris EH, Sato R;

XX DR WPI; 1997-132630/12.

XX PT Chlamydomonas reinhardtii derived DNA conferring resistance to
XX PT porphyrin herbicides, used to transform plant or algal cells to
XX PT confer selective herbicide resistance

XX PS Claim 14; Page 40-41; 50pp; English.

XX CC The present sequence represents an isolated, purified DNA fragment
XX CC that confers resistance to porphyrin-accumulating type herbicides
XX CC when expressed in plant or algal cells. The DNA fragment can be used
XX CC in plants and algae to confer resistance to herbicides such as
XX CC N-(4-chloro-2-fluoro-5-propargyloxy) phenyl-3,4,5,6
XX CC tetrahydrothiophene, oxadiazon and oxyfluorfen. It permits the use
XX CC of porphyrin herbicides during crop cultivation and thus facilitates
XX CC weed control in crops. The gene can also be used as a resistance-type
XX CC genetic marker in plant genetic engineering.

XX SQ Sequence 3381 BP; 670 A; 1008 C; 1058 G; 645 T; 0 other;

Query Match 96.6%; Score 136.2; DB 18; Length 3381;
Best Local Similarity 97.9%; Pred. No. 1.6e-21;
Matches 138; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gccgcgagccctgggctcttcgactaccgcgagcggtggcgccgctgacgtgctgac 60
|||||
Db 3127 gccgcgagccctgggctcttcgactaccgcgagcggtggcgccgctgacgtgctgac 3068
|||||
Qy 61 ccgctgagcgcgtgagcgagcgcaaggcctcgacgagcggtgctgcccggctcggt 120
|||||
Db 3067 ccgctgagcgcgtgagcgagcgcaaggcctcgacgagcggtgctgcccggctcggt 3008
|||||
Qy 121 cagctgacccgcgcacgacg 141
|||||
Db 3007 cagctgacccgcgcacgacg 2987
|||||

RESULT 7

T60965/c
ID T60965 standard; DNA; 3383 BP.

XX AC T60965;

XX DT 14-MAY-1997 (first entry)

XX DE Porphyrin-accumulating type herbicide resistance gene.

XX KW Herbicide resistance; porphyrin; transgenic plant; alga;
XX KW weed control; selectable marker; ds.

XX OS Chlamydomonas reinhardtii strain RS-3.

XX PN W09704088-A1.

XX PD 06-FEB-1997.

XX PF 20-JUL-1995; 95WO-US09098.

XX PR 20-JUL-1995; 95WO-US09098.

XX XX (SUMO) SUMITOMO CHEM CO LTD.

XX PA (UYDU-) UNIV DUKE.

XX PI Boynton J, Gillham NW, Harris EH, Sato R;

XX DR WPI; 1997-132629/12.

XX PT DNA conferring resistance to porphyrin herbicides - isolated from
XX PT Chlamydomonas reinhardtii, useful to transform plant or algal cells
XX PT to confer selective herbicide resistance

XX PS Claim 14; Page 39-41; 52pp; English.

XX CC A porphyrin accumulating-type herbicide resistance gene (T60965)
XX CC was isolated from Chlamydomonas reinhardtii strain RS-3. Sau3AI
XX CC fragments of genomic DNA were ligated into SuperCos-1 vector and
XX CC used to transform E. coli. The DNA was extracted and used to
XX CC transform the normally sensitive strain CC48 to provide herbicide
XX CC resistance to N-(4-chloro-2-fluoro-5-propargyloxy)phenyl-3,4,5,6-
XX CC tetrahydrothiophene. The DNA was then analysed to provide the
XX CC porphyrin herbicide resistance gene. The gene can be incorporated
XX CC into a plasmid and used to confer resistance to porphyrin
XX CC accumulating-type herbicides in plant or algal cells. It can also
XX CC be used as a genetic resistance marker.

XX SQ Sequence 3383 BP; 671 A; 1008 C; 1057 G; 647 T; 0 other;

Query Match 96.6%; Score 136.2; DB 18; Length 3383;
Best Local Similarity 97.9%; Pred. No. 1.6e-21;
Matches 138; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CC arachnids, nematodes and insects. Biosynthesis of spinosyns occurs via
CC stepwise condensation and modification of carboxylic acid precursors
CC generating a linear polyketide which is modified further. The DNA
CC sequence 221501 contains 24 genes and open reading frames involved in
CC spinosyn biosynthesis, Y39297-Y39319. The genes identified in S. spinosa
CC as having a role in the production of spinosyns are useful to improve
CC yields of spinosyns, and for creating new spinosyns e.g. by mutagenesis,
CC or interruption of steps in spinosyn biosynthesis. The modified spinosyns
CC may be a new insect control agent or serve as substrates for further
CC chemical modification and the creation of new semi-synthetic spinosyns.
CC The genes are also useful to isolate similar sequences from S. spinosa or
CC other species by hybridization.
XX
SQ Sequence 2310 BP; 356 A; 759 C; 853 G; 342 T; 0 other;

Query Match 27.0%; Score 38; DB 20; Length 2310;
Best Local Similarity 57.6%; Pred. No. 1.8;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Oy 1 gcgcgagccctggcctcttcgactaccgcgcgtggcgccgtgacgtgtgtgtac 60

Db 1213 gtcgcgacgcgttgggtctcttcgagacgcggaagacgcgagctgcgaccggtg 1272

Oy 61 ccgctgagcgcgtgcggagagcgcaagcgcctcgcgagcgcgtgcgcggcttcg 118

Db 1273 gtgatacgcgcgcgtgtacacgcgcgtggacgcgcgcgagtcgcgaccgcgaccgcg 1330

RESULT 12

V84525/c
ID V84525 standard; DNA; 1288 BP.

XX AC V84525;

XX AC V84525;

XX DT 01-MAR-1999 (first entry)

XX DE Human secreted protein gene 115 clone HDTA95.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX OS Homo sapiens.

XX OS WO9854963-A2.

XX PD 10-DEC-1998.

XX PF 04-JUN-1998; 98WO-US11422.

XX PR 18-DEC-1997; 97US-0070923.

XX PR 06-JUN-1997; 97US-0048877.

XX PR 06-JUN-1997; 97US-0048881.

XX PR 06-JUN-1997; 97US-0048884.

XX PR 06-JUN-1997; 97US-0048893.

XX PR 06-JUN-1997; 97US-0048896.

XX PR 06-JUN-1997; 97US-0048899.

XX PR 06-JUN-1997; 97US-0048915.

XX PR 06-JUN-1997; 97US-0048949.

XX PR 06-JUN-1997; 97US-0048964.

XX PR 06-JUN-1997; 97US-0049020.

XX PR 06-JUN-1997; 97US-0049375.

XX PR 05-SEP-1997; 97US-0057628.

XX PR 05-SEP-1997; 97US-0057635.

XX PR 05-SEP-1997; 97US-0057644.

XX PR 05-SEP-1997; 97US-0057647.

PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057661.
PR 05-SEP-1997; 97US-0057667.
PR 05-SEP-1997; 97US-0057761.
PR 05-SEP-1997; 97US-0057784.
PR 05-SEP-1997; 97US-0057770.
PR 05-SEP-1997; 97US-0057775.
PR 05-SEP-1997; 97US-0057778.
PR 06-JUN-1997; 97US-0048875.
PR 06-JUN-1997; 97US-0048878.
PR 06-JUN-1997; 97US-0048882.
PR 06-JUN-1997; 97US-0048885.
PR 06-JUN-1997; 97US-0048894.
PR 06-JUN-1997; 97US-0048897.
PR 06-JUN-1997; 97US-0048900.
PR 06-JUN-1997; 97US-0048916.
PR 06-JUN-1997; 97US-0048962.
PR 06-JUN-1997; 97US-0048970.
PR 06-JUN-1997; 97US-0048974.
PR 06-JUN-1997; 97US-0049373.
PR 05-SEP-1997; 97US-0057584.
PR 05-SEP-1997; 97US-0057629.
PR 05-SEP-1997; 97US-0057642.
PR 05-SEP-1997; 97US-0057645.
PR 05-SEP-1997; 97US-0057648.
PR 05-SEP-1997; 97US-0057651.
PR 05-SEP-1997; 97US-0057662.
PR 05-SEP-1997; 97US-0057668.
PR 05-SEP-1997; 97US-0057762.
PR 05-SEP-1997; 97US-0057765.
PR 05-SEP-1997; 97US-0057771.
PR 05-SEP-1997; 97US-0057776.
PR 06-JUN-1997; 97US-0048876.
PR 06-JUN-1997; 97US-0048880.
PR 06-JUN-1997; 97US-0048883.
PR 06-JUN-1997; 97US-0048892.
PR 06-JUN-1997; 97US-0048895.
PR 06-JUN-1997; 97US-0048898.
PR 06-JUN-1997; 97US-0048901.
PR 06-JUN-1997; 97US-0048917.
PR 06-JUN-1997; 97US-0048963.
PR 06-JUN-1997; 97US-0048971.
PR 06-JUN-1997; 97US-0049019.
PR 06-JUN-1997; 97US-0049374.
PR 05-SEP-1997; 97US-0057627.
PR 05-SEP-1997; 97US-0057634.
PR 05-SEP-1997; 97US-0057643.
PR 05-SEP-1997; 97US-0057646.
PR 05-SEP-1997; 97US-0057649.
PR 05-SEP-1997; 97US-0057654.
PR 05-SEP-1997; 97US-0057666.
PR 05-SEP-1997; 97US-0057760.
PR 05-SEP-1997; 97US-0057763.
PR 05-SEP-1997; 97US-0057769.
PR 05-SEP-1997; 97US-0057774.
PR 05-SEP-1997; 97US-0057777.

(HUMA-) HUMAN GENOME SCI INC.

PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA;

PI Fan P, Feng P, Ferrie AM, Fischer CL, Florence C;

PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW;

PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;

XX Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;

XX WPI; 1999-059865/05.

XX P-PSDB; W88648, W88934.

XX New isolated human genes and the secreted polypeptides they encode -

XX useful for diagnosis and treatment of e.g. cancers, neurological

XX disorders, immune diseases, inflammation or blood disorders

XX Claim 4; Page 377-378; 772pp; English.

XX The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W88534 to W88756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents a gene encoding a human secreted protein
CC (see descriptor line for gene number and clone identification).
XX
XX Sequence 1288 BP; 356 A; 264 C; 295 G; 370 T; 3 other;

Query Match 25.4%; Score 35.8; DB 20; Length 1288;
Best Local Similarity 53.3%; Pred. No. 5.7;
Matches 73; Conservative 1; Mismatches 63; Indels 0; Gaps 0;
QY 1 gccgcgagccctgggtctcttcactaccgcgcggtggcgccgtgacgtgtctgtac 60
DB 153 ggcggcgccggccctgggggtcgatcgctccggctgcggcgagcgagcgagcgcg 94
QY 61 ccgctgagccctgcgggagagcgcaagcctcgagcgggtccgctgcggggttcggt 120
DB 93 gaggagaccgagagaggaacgtgttcagctgtgctccgcgcgcgtccgagcgccgcg 34
QY 121 cagctgaccccgagc 137
DB 33 agcgctgcatcaatgcgcg 17

RESULT 13
A63462
ID A63462 standard; DNA; 3035 BP.
XX
AC A63462;
XX
DT 06-MAR-2001 (first entry)
XX
DE Streptomyces globisporus C-1027 gene cluster sgca and sgcb genes.
XX
KW Eneidyne C-1027 biosynthesis gene cluster; apoprotein; chromophore;
KW sgca; sgcb; cancer; ds.
XX
OS Streptomyces globisporus.
XX
FH Key Location/Qualifiers
FT CDS 101..1099
FT /*tag= a
FT /product= "Sgca dndp-glucose 4,6-dehydratase"
FT CDS 1144..2408
FT /*tag= b
FT /product= "Sgcb transmembrane efflux protein"
XX
PN WO200040596-A1.
XX
PD 13-JUN-2000.

XX 06-JAN-2000; 2000WO-US000446.
PF
XX
XX 06-JAN-1999; 99US-0115434.
PR
XX 05-JAN-2000; 2000US-0477962.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Shen B, Liu W, Christenson SD, Standage S;
PI WPI; 2000-465947/40.
XX
XX P-PSDB; B13604, B13605.
DR
XX
XX Isolated nucleic acid comprising a nucleic acid encoding any of C-1027
PT open reading frames (ORFs) -7 to 42, excluding ORF 9 (caga), useful for
PT the production of enediyne C-1027 antitumour antibiotics -
XX
XX Example 1; Fig 6; 160pp; English.
XX
XX The present sequence is the sgca and sgcb genes which form part of the
CC enediyne C-1027 gene cluster in Streptomyces globisporus. Sgca encodes a
CC dndp-glucose 4,6dehydratase enzyme, and sgcb encodes a transmembrane
CC efflux protein. Eneidyne C-1027 is an antibiotic, consisting of an
CC apoprotein and a non-peptidic chromophore, which acts by damaging DNA.
CC The sequences within the gene cluster, and the proteins they encode, can
CC be used in the treatment of cancer, along with antagonists of the
CC protein.
XX
XX Sequence 3035 BP; 440 A; 1101 C; 1029 G; 465 T; 0 other;

Query Match 25.1%; Score 35.4; DB 21; Length 3035;
Best Local Similarity 54.1%; Pred. No. 6.4;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 1 gccgcgagccctgggtctcttcactaccgcgcggtggcgccgtgacgtgtctgtac 60
DB 977 ggctacgagccctgggtctcttcactggagcgccctggcgagcggtgaagtgtgtaccac 1036
QY 61 ccgctgagcgccctgcgggagagcgcaagcctcgagcgggtccgctgcggggttcggt 120
DB 1037 gagaaccttcgtgtgtgagcgcgtgaaggaagcgccctcctcgagcgccgcgcgc 1096
QY 121 cagctgaccccg 133
DB 1097 tgacgagcagccac 1109

RESULT 14
A63349
ID A63349 standard; DNA; 42000 BP.
XX
AC A63349;
XX
DT 06-MAR-2001 (first entry)
XX
DE Streptomyces globisporus C-1027 gene cluster ORF -7-25.
XX
KW Eneidyne C-1027 biosynthesis gene cluster; apoprotein; chromophore;
KW cancer; ds.
XX
OS Streptomyces globisporus.
XX
FH Key Location/Qualifiers
FT CDS complement (8..658)
FT /*tag= a
FT /product= "ORF -7 protein"
FT CDS complement (930..1478)
FT /*tag= b
FT /product= "ORF -6 protein"
FT CDS complement (1649..2713)
FT /*tag= c
FT /product= "ORF -5 protein"


```
FT      /*tag= ap
FT      /product= "ORF 34 protein"
FT      56026..56880
FT      /*tag= aq
FT      /product= "proline oxidase"
FT      complement (56927..57736)
FT      /*tag= ar
FT      /product= "ORF 36 protein"
FT      complement (57833..58304)
FT      /*tag= as
FT      /product= "ORF 37 protein"
FT      complement (58440..60095)
FT      /*tag= at
FT      /product= "P450 hydroxylase"
FT      complement (60092..60621)
FT      /*tag= au
FT      /product= "ORF 39 protein"
FT      60940..62019
FT      /*tag= av
FT      /product= "ORF 40 protein"
FT      62045..62899
FT      /*tag= aw
FT      /product= "ORF 41 protein"
FT      complement (62787..63164)
FT      /*tag= ax
FT      /product= "ORF 42 protein"
```

WO200040596-A1.

13-JUL-2000.

06-JAN-2000; 2000WO-US00446.

06-JAN-1999; 99US-0115434.

05-JAN-2000; 2000US-0477962.

(REGC) UNIV CALIFORNIA.

Shen B, Liu W, Christenson SD, Standage S;

WPI: 2000-465947/40.

DR P-PSDB; B13554, B13555, B13556, B13557, B13558, B13559, B13560, B13561,
B13562, B13563, B13564, B13565, B13566, B13567, B13568, B13569, B13570,
B13571, B13572, B13573, B13574, B13575, B13576, B13577, B13578, B13579,
B13580, B13581, B13582, B13583, B13584, B13585, B13586, B13587, B13588,
B13589, B13590, B13591, B13592, B13593, B13594, B13595, B13596, B13597,
B13598, B13600, B13601, B13602, B13603, B13606, B13607.

XX Isolated nucleic acid comprising a nucleic acid encoding any of C-1027
PT open reading frames (ORFs) -7 to 42, excluding ORF 9 (cagA), useful for
PT the production of enediyne C-1027 antitumour antibiotics -

Claim 1; Page 78-157; 160pp; English.

XX The present sequence is the enediyne C-1027 gene cluster from
CC Streptomyces globisporus. Enediyne C-1027 is an antibiotic, consisting of
CC an apoprotein and a non-peptidic chromophore, which acts by damaging DNA.
CC The sequences within the gene cluster, and the proteins they encode, can
CC be used in the treatment of cancer, along with antagonists of the
CC protein. Each of the open reading frames is specifically claimed,

Query Match 25.1%; Score 35.4; DB 21; Length 63164;
Best Local Similarity 54.1%; Pred. No. 4.9;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 1 gccgcgagccctcgtccttcgactaccgcgcgctggcgccgtgacgctgtcgtac 60

Db 29466 ggctacgagccctcgtcgtccttcgagcgcgcgctggcgccgtggaagtgtaccac 29525

QY 61 ccgctgagccgctgcgggagagcgcaagcctcgacgggtccgtgcgggcttcggt 120

Db 29526 gagaaccgttcgtggtgcgcgcgtgaaggagcgcgcgctcctctggacgcgcgtcggc 29585

QY 121 cagctgcacccgc 133
Db 29586 tgacgcagccac 29598

Search completed: June 2, 2001, 00:43:34
Job time: 2322 sec

